

# THIS WEEK

## EDITORIALS

**CHURCH** Scientists should welcome the election of Pope Francis I **p.282**

**WORLD VIEW** New head sought for European Research Council **p.283**

**LAB SAFETY** Plastic pills protect DNA from radiation and heat **p.285**



## Wasted energy

*The burning off of gas during oil extraction is environmentally unsound and unjustifiable. The United States should instead be seeking to make use of this natural resource.*

One could perhaps forgive the oilmen of the past. In their pursuit of black gold, they simply burnt off the natural gas that was extracted from the rocks alongside their precious oil. It was a time when Earth's bounty seemed to expand without consequence in the face of human ingenuity and technological prowess. The incentive to invest in the infrastructure to capture the gas and bring it to market simply did not exist.

But that is no longer true. There is no justification for the large-scale burning off, or flaring, of natural gas by today's oil industry, particularly in the United States, which is home to the most mature and advanced oil and gas industry in the world. Nearly one-third of the raw gas that is pumped out of the Bakken shale formation in North Dakota — a prime target of the new hydrofracturing and horizontal drilling technologies — is burned *in situ*.

True, flaring is preferable to venting gases such as methane, butane and propane directly into the atmosphere, but it still has a detrimental effect on both the global climate and the local air quality. And because companies are exempt from paying taxes or royalties on vented gas for the first year, at least, it is also bad for the public purse. As a result, the public gets a smaller return on the environmental price being paid to recover this oil — the inevitable impacts on public infrastructure, air and water resources, and on the landscape itself.

At a conservative estimate, this North Dakota flaring meant some 3.9 million tonnes of carbon dioxide were emitted last year, the equivalent of the annual emissions from 750,000 vehicles. Worse, research into flaring has begun to find evidence of potentially widespread methane leakage from shale operations, if not outright venting of the gas (see page 290). Methane is a powerful greenhouse gas, so the environmental price is likely to be even higher.

There are solutions. In North Dakota, the state could halt the practice of flaring except when necessary for safety reasons, or it could discourage companies from flaring by making them pay regular taxes and royalties on flared gas. This could delay the development of shale deposits, but that might be a good thing because it would give landowners and government agencies more time to work out how to regulate the environmental and social challenges that accompany energy booms such as shale exploitation. The companies have plenty of motivation to get shale resources out of the ground, and there can be little doubt that they would find ways to exploit the gas currently being flared, perhaps by exporting it. And methane emissions could be better controlled if the US Environmental Protection Agency regulated it as a greenhouse gas and instituted stricter rules across the oil and gas industry.

The US shale boom has been a boon to the struggling economy, providing jobs and government revenue in many far-flung places. The resulting oil production has allowed the United States to reduce foreign imports, and the plentiful shale-gas resources have lowered demand for coal, thereby curbing greenhouse-gas emissions in the power sector. But it will be up to scientists to pin down the full suite of impacts from

the new oil and gas developments and to help policy-makers better understand the choices that they are making.

On 15 March, President Barack Obama proposed creating a US\$2-billion Energy Security Trust to advance research and development into low-carbon transport alternatives. It would be funded over a decade by diverting a portion of the proceeds from federal oil and gas development, which are poised to grow thanks to the shale bonanza in North Dakota and beyond. It is a good idea as far as it goes, although once again it is hard not to despair at the general lack of ambition on climate issues in Washington. Improbable as it may be, a federal

carbon tax would raise more money and would send an important signal to the energy industry that it needs to control its greenhouse-gas emissions.

The US economy is already benefiting from shale developments, and the country might even be able to lock in a one-time emissions reduction as part of a broader shift from coal to cleaner-burning natural gas in the coming years. A logical part of that equation is to kill off the current fashion for flares. ■

***"The public gets a smaller return on the environmental price being paid to recover this oil."***

## CITES for sore eyes

*Successes at last week's wildlife-trade treaty meeting must be backed up with action.*

After the final votes were cast at the 16th conference of parties to the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) last week, it became clear that something remarkable had happened.

Preliminary decisions earlier in the meeting, which ran from 3 to 14 March in Bangkok, had increased protection for a plethora of species and given cautious hope to zoologists and botanists. The final vote, by representatives of 170 of the states that are signed up to the treaty, rubber-stamped those decisions and sent researchers and conservationists out with broad smiles.

The conference of the parties to CITES deserves praise and recognition for placing a number of species of sharks and rays onto its Appendix II, which regulates trade in animals that are not deemed to be at immediate risk of extinction.

CITES delegates seem to have discovered an overdue and very welcome willingness to step on the toes of influential commercial interests that consistently oppose such restrictions, as they moved to protect

marine species of huge commercial importance. They also saw the value in the wood from the trees: several species of tropical hardwoods were added to Appendix II.

Key agreements on trade in elephant ivory were strengthened, to specify the need for campaigns to reduce demand. This is the best — and possibly the only — way to save elephant populations in the long term, say many researchers who study the illegal trade that takes poached ivory from Africa to markets in countries such as China and Thailand.

There are already promising signs that public awareness of the threat to elephants is growing in China, as demonstrated by the campaigning of one of the country's biggest (in more ways than one) celebrities, the basketball star Yao Ming. Similar campaigns against the trade in shark fins — whose status as a delicacy in some Asian countries is often blamed for declines in shark numbers — seem also to be finding a receptive audience.

At the CITES meeting, there was also success for attempts to clamp down on the ivory trade through increased forensic scrutiny of seized ivory and the stockpiles of tusks kept in many African nations. This is a significant victory. Such DNA analysis should provide crucial information on the illegal trade and open up new avenues to combat it (see *Nature* 494, 411–412; 2013). It is also a validation of the hard work and campaigning put into this problem. The solid evidence base produced by the CITES projects Monitoring the Illegal Killing of Elephants and the Elephant Trade Information System has focused global attention on the resurging crisis of elephant slaughter.

These successes carry important lessons. Popular and scientific reports of the threats to elephants, sharks and other species have helped to tip the political balance in favour of strengthening regulations. Non-governmental organizations deserve credit for raising the

alarm, as do politicians for heeding the warnings.

Let us not get carried away. There were also disappointments at CITES — notably the failure to stop trade in polar bears and their parts. Some conservationists also wanted even tougher moves to clamp down on elephant and rhino poaching, including trade sanctions, which were rejected. Overall, however, the post-meeting mood was jubilant — and rightly so.

Serious questions must now be asked about the positive part that CITES can play in future marine conservation. If CITES wants to make progress in this sphere, then it must bring pressure to bear on special-

***“Delegates seem to have discovered a willingness to step on the toes of influential commercial interests.”***

ist fisheries-management bodies, many of which have attracted criticism for allowing species such as tuna to be fished beyond sustainable limits. Those bodies have three years to put their houses in order before the next CITES meeting, or CITES will be obliged to do it for them. (The convention does not have a perfect record here, however — it failed in efforts to protect tuna populations in 2010, and dodged the issue at the latest meeting.)

The decisions passed last week will not by themselves save a single animal or plant. Proper monitoring is essential. To build on the successes of the meeting, funders must provide stable financing for continued research on the welfare of those species that now fall under the protection of CITES — and those that do not.

CITES took a great step forward last week. Its success should inspire all those who push for evidence-based policy. Perhaps most importantly, it shows that international meetings that seethe with dissonant agendas and actors are not always toothless talking shops. ■

# A pope for today

*Latest pontiff looks to enhance social relevance of Catholic Church.*

Whether or not you are a believer, it is hard not to like the man. In the few days since the white smoke began to billow from the Sistine Chapel in Vatican City, the world has learned a little about Jorge Mario Bergoglio, the 76-year-old elected as Pope Francis I. The first pope from Latin America, as archbishop of Buenos Aires he eschewed the trappings of the office, forgoing a mansion for a small apartment, preferring to take the bus than use a chauffeur, and dedicating himself to pastoral work in the slums. The affable Pope Francis has also already wooed the public (and much of a fawning media) with his disarming humility and common touch — and his obvious flair for ad-libbing and humour. It is clear that Francis's papacy marks a break with the past, a new distinctive and refreshing papal style, and an ambition to focus on social relevance and justice. “How I would like a Church which is poor and for the poor!” he said.

We also learnt that the man obtained his first degree in chemistry, a later one in philosophy and another in theology, and that he has taught literature and psychology at universities. That broad education, academic bent and humility are hardly a surprise because Bergoglio is the first Jesuit pope. The Jesuits, the largest order in the Catholic Church, are its intellectual elite and known for their independent thinking. They also vow to live lives of austerity and never to seek high office in the Church — let alone pope. They have focused on issues of social and economic injustice, and less on doctrine than do career clergy. They have long worked as missionaries, and are probably best known for their creation and running of some of the world's top schools and universities. Many are also scientists.

We know little about Bergoglio's views on scientific issues, which he has hardly written about. The hordes of scientists among the Church's 1.2 billion baptized members would like to hear more. And his chemistry degree in itself says little about the Pope's attitudes to science. But what is clear is that, contrary to widespread belief, the modern Catholic Church is science-friendly and Pope Francis will no doubt continue, and perhaps deepen, that tradition. The Church's strong support for Darwinian evolution, for example, contrasts sharply with the backwards unscientific belief in creationism of many US evangelicals and lawmakers — a concept that Pope Benedict XVI rightly criticized in 2007 as “absurd”. Priests also gave us Mendelian genetics and contributed to the theory of the Big Bang.

Moreover, recent popes have substantially increased efforts to engage in dialogue with scientists on a host of issues, from embryonic stem-cell research and genetically modified crops to *in vitro* fertilization, abortion and euthanasia — and in the future will no doubt increasingly do so on advances in neuroscience and genetics, including prenatal screening. Scientists who have taken part in such discussions tell of thought-provoking and constructive debates, with the Church being open to ideas and often changing doctrines as a result. A damaging exception is its long-held opposition to the use of condoms to prevent the spread of HIV, and it can only be hoped that Pope Francis will have a more enlightened approach.

But whereas doctrines can be tweaked, the Church will not compromise on its central dogmas, such as the sanctity of human life and that life begins at conception. Science and faith can provide complementary world views, with progress in science informing and often challenging the rationale of Church doctrines, and vice versa: faith can often add much-needed dimensions of ethics and social justice to advances in science and their impact on society. Clashes are inevitable between people of different beliefs, but both science and religion are best served by building bridges across the divides. How Pope Francis responds to issues where the two meet will be an important mark of the man. ■

➔ **NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunjv](http://go.nature.com/xhunjv)



DFG



## A chance to drive forward Europe's science

*The new head of the European Research Council will have more power and focus than former presidents, says Ernst-Ludwig Winnacker.*

**W**ANTED: An ambitious and skilled science administrator to oversee a continent-wide research programme with considerable impact and a steadily increasing budget. Diplomatic skills essential. Must be able to travel.

One of the most appealing jobs in global science is now available.

The European Research Council (ERC) is looking for a new president, and the European Commission, which guarantees the autonomy and integrity of the council, has set up an independent search committee to find just the right person.

What makes the job so attractive? The answer is not too difficult. The ERC is a resounding success. Since its launch more than six years ago, it has supported more than 3,000 researchers. It is the jewel in the crown of European Union (EU) research.

A decade ago, things were quite different. European scientists, many of whom hailed from countries that did not have national funding agencies, were fighting for a Europe-wide agency that could promote excellence in basic research, and could match what they saw in the United States.

I was involved in that effort, which was universally seen as a good idea, but difficult to get going. Early, cautious efforts aimed to create an 'agency of the agencies'. Only when the European Commission got involved did the concept begin to float.

Because the European Commission will remain in the financial driving seat for the coming years, the most challenging part of the new job will be to work with the body. Commission-bashing is easy and fashionable, but it is usually superficial. The complexity of the commission's operations is daunting — to understand and accept that will require enormous patience, knowledge and communication skills.

Change is never rapid, and an evolutionary rather than a revolutionary type of personality may therefore be required. The commission and its representatives hold the ERC in high esteem, and they have made an impressive effort to reduce its administrative burdens. Nevertheless, there remains some way to go.

The first two presidents of the ERC — Fotis Kafatos and Helga Nowotny — did an excellent job. And thanks to the ingenuity of the Scientific Council — the ERC's strategy and decision-making body — and the strong commitment of the administrative staff, the ERC works well.

This is not to say that the third president will be able to climb into a neatly made-up bed. To raise its attractiveness, the job has been redesigned to be more powerful and focused than for the former presidents.

The restructuring follows the 2011 recommendation of a task force that investigated the ERC's governance. The positions of president of the Scientific Council and of secretary general have been fused to create a President *sui generis*,

THE JOB HAS BEEN  
**REDESIGNED**  
TO BE  
**MORE**  
**POWERFUL**  
AND FOCUSED.

based in Brussels and thus able to deal directly on site with the various day-to-day tussles of the job.

Still, president number three will face several challenges. One is the extremely uneven distribution of returns on investment to the various member states and associated countries. Strict adherence to the council's core principle of funding excellence alone has highlighted extreme differences in scientific competence over the European continent. More than 98% of the ERC's grants are awarded to scientists in the old EU-15, with mere crumbs thrown to newer member countries. Although many of these countries have fine academic traditions, their supporting infrastructures tended to be neglected during the time of the iron curtain, and may take a generation or two to restore. The ERC cannot remedy this deficiency alone, but it could have an efficient advisory role.

The ERC, like much of science, still has a gender problem. For example, 29% of applicants for a starting grant and 15% of applicants for advanced grants are women, yet the success rates for female scientists are consistently at least 2% lower than for men. The ERC is, of course, aware of this, and of the unconscious bias that may still be affecting peer review. The ERC must not only remain sensitive to the issue, but also find a more proactive way to tackle it.

A particularly interesting problem is the evolution of the relationship between the ERC and the national funding councils. Some in Europe think that the national councils should dissolve and only the ERC survive. But mono-cultures are never ideal and can even be detrimental.

It would be in the best interest of both the ERC and the national research councils to develop plans for continued coexistence. A first step could be to give applicants from new member states access to the financial resources of other ERC nations. Bulgarian or Estonian scientists should be able to win money from the, say, German, Dutch or UK research councils, even if working in their home countries. Why should the rich countries participate? Because it would broaden the base of their own competitions and because it could rapidly improve the opportunities for scientists in the new member states to raise ERC grants, thereby reducing the unhealthy divide between East and West.

Finally, and perhaps most importantly, the financial outlook for the ERC is positive. Its budget for 2014–20 will rise significantly. Not even our colleagues in the United States can say that.

The ERC needs a president with a truly European mindset, and one who believes that it is important to build a broader base for frontier research. I hope our best and brightest will take note. ■

**Ernst-Ludwig Winnacker** was the first secretary general of the ERC and is now secretary general of the Human Frontier Science Program Organization in Strasbourg, France.  
e-mail: [elwinnacker@gmail.com](mailto:elwinnacker@gmail.com)

➔ **NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/ht1dbd](http://go.nature.com/ht1dbd)

# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## QUANTUM COMPUTING

### Noisy quantum calculations

Researchers have proposed a way to make use of noise in quantum computing.

Isolating systems from 'noisy' environmental interactions has been the focus of much research. Jens Eisert and his colleagues at the Free University of Berlin, however, propose an alternate framework in which information is encoded in a quantum system combined with its noisy environment.

To do this, the authors exploit an effect called the cutoff phenomenon, in which combined systems suddenly change states. This phenomenon defines points in time to launch calculations or read out results.

*Phys. Rev. Lett.* 110, 110501 (2013)

## MEDICINE

### Blood test for cancer DNA

The DNA that is shed into the bloodstream from dying cancer cells might one day be used to monitor the disease.

Carlos Caldas at the Cancer Research UK Cambridge Institute and his colleagues identified genetic markers to track disease in 30 women being treated for advanced breast cancer. They found that a technique to detect these markers in cancer-cell DNA in the blood was more sensitive than tests that look for circulating tumour cells or for a cancer antigen. The amount of tumour DNA also roughly correlated with the response to treatment.

The authors suggest that blood tests based on tumour DNA could be quicker and less-invasive than biopsies.

*N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1213261> (2013)

For a longer story on this research, see [go.nature.com/luwx3m](http://go.nature.com/luwx3m)

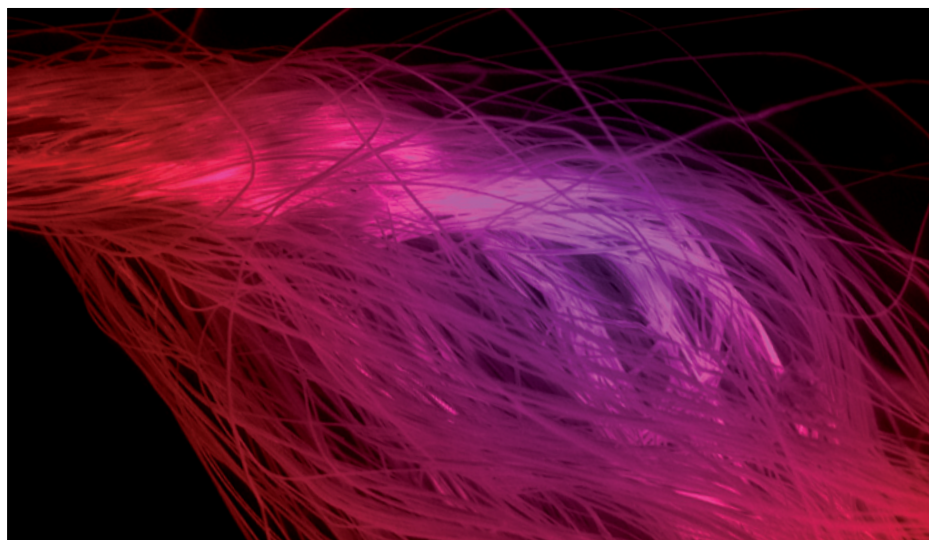
## IMMUNOLOGY

### Flu vaccines may improve with age

A subset of white blood cells may determine who is best protected by seasonal flu vaccines.

Hideki Ueno at Baylor Research Institute in Dallas, Texas, and Octavio Ramilo at the Nationwide Children's Hospital in Columbus,

Ohio, and their colleagues injected 49 adults and 19 children with flu vaccine. The researchers tracked cell types in individuals over time and discovered a particular class of immune cells that is activated on vaccination. These cells can boost production of existing antibodies that fight flu, but they do not induce production of new antibodies. This suggests that the current flu-vaccine strategy may be less effective in young children and against rare types of flu, such as H5N1. *Sci. Transl. Med.* 5, 176ra32 (2013)



## NANOMATERIALS

### Fibres toughen when stretched

When most tough fibres are stretched to make them thinner, they become brittle. But a group led by Yuris Dzenis at the University of Nebraska–Lincoln has shown that this is not always the case.

The researchers made polyacrylonitrile fibres (pictured) using a technique called electrospinning. As their diameters narrowed to below 250 nanometres, the fibres became tougher and so were less prone to fracture, but

did not lose their strength. Nanofibres were up to 10-fold tougher and stronger than the best commercial fibres.

Dzenis suggests that the toughening is possible because the nanofibres are less crystalline than larger fibres. He thinks that the fibres could be used in load-bearing aerospace structures and bulletproof materials.

*ACS Nano* <http://dx.doi.org/10.1021/nn400028p> (2013)

## MICROBIOLOGY

### Life on the seabed, and below

The dark abyss of the Mariana Trench in the western Pacific Ocean hosts a surprisingly active bacterial community.

By using an automated deep-sea instrument that they designed, a team led by Ronnie Glud at the University of Southern Denmark in Odense measured rates of biological oxygen consumption in the almost 11-kilometre-deep ocean trench. The high rates they found there suggest that degradable organic matter

DIMITRY PAKOV/JOEL BREHM/YURIS DZENIS

is being deposited at the bottom of the trench, allowing microbes to thrive in this secluded habitat.

A second team reports that life exists in an even more unlikely environment. While at the University of North Carolina at Chapel Hill, Mark Lever and his colleagues found hydrogen-loving microbes in the oceanic crust, deep beneath the seabed off the west coast of North America. An incubation experiment over several years showed that the organisms can derive energy from geochemical reactions between iron compounds and the sea water that permeates through cracks in the rocky crust.

*Nature Geosci.* <http://dx.doi.org/10.1038/ngeo1773> (2013);

*Science* 339, 1305–1308 (2013)

For a longer story on this research, see [go.nature.com/ppzjaz](http://go.nature.com/ppzjaz)

## CHEMISTRY

## Protecting DNA in silica ‘amber’

High temperatures and harsh chemicals degrade DNA, but not if it is protected by a skin of silica.

Robert Grass and his colleagues from the Swiss Federal Institute of Technology in Zurich immobilized DNA on charged silica particles and grew a 10-nanometre silica layer on top. The encapsulated DNA was variously subjected to strong ultraviolet radiation, humid 120°C heat and attack by reactive chemicals. Each time, the molecule could be recovered by dissolving the silica with a weak acid. The DNA remained largely undamaged (much as fossilized DNA is protected in amber) and its sequence could be easily read. The researchers used a 195°C press to create plastic pills (pictured) containing the protected DNA, which



PHILLIP STOSSEL

they suggest could be added to consumer goods as high-security barcodes.

*Angew. Chem.* <http://dx.doi.org/10.1002/anie.201208135> (2013)

## BIOENGINEERING

## Mobile worm microscope

Smart phones may offer an affordable, portable way to diagnose intestinal worm infection, the most common infection in developing countries.

Isaac Bogoch at Toronto General Hospital in Canada and his group transformed an iPhone into a microscope by attaching an inexpensive lens to its camera. Almost 200 stool samples from children in Pemba Island, Tanzania, were mounted on slides and studied using the microscope. Experts identified more than 60% of worm-free slides, and nearly 70% of samples with eggs — of these, eggs from the large roundworm *Ascaris lumbricoides* had the highest rate of detection, at 81%.

Although similar portable microscopes have been made before, most have stayed in the laboratory. The authors say that this is the first field-based clinical test of such devices for intestinal worms, and they predict that higher-resolution imaging will boost the accuracy of smart-phone diagnoses.

*Am. J. Trop. Med. Hyg.* <http://dx.doi.org/10.4269/ajtmh.12-0742> (2013)

## GEOLOGY

## Quakes shake the gold out

Earthquake after earthquake may build up economically important gold deposits by concentrating the precious metal along quartz-rich veins under the ground.

Dion Weatherley of the University of Queensland in Brisbane, Australia, and Richard Henley at the Australian National University in Canberra modelled jogs, or

## COMMUNITY CHOICE

The most viewed papers in science

## MEDICAL RESEARCH

## Mice poor models for inflammation

**HIGHLY READ**  
on [www.pnas.org](http://www.pnas.org)  
in February

In a systematic evaluation of how well mice mimic human inflammatory responses, the popular disease model receives poor marks.

A large multi-institute group co-led by Ronald Tompkins of Harvard Medical School in Boston, Massachusetts, studied genome-wide gene expression in blood samples from 35 healthy people and 411 people with burns, severe blunt trauma or sepsis. They also used public databases to gather human data for other acute inflammatory stresses. Shifts in gene expression as a result of stress were highly similar across the different groups of patients, but correlated poorly with those in a mouse model. Such disparities could lead to misleading results in studies of disease or in the development of therapies, the authors say. The quality of a disease model should be assessed not just for the appearance of disease but also for its biomolecular fidelity, they suggest.

*Proc. Natl Acad. Sci.* 110, 3507–3512 (2013)

cracks, that cross geological faults. Ground motions can drastically increase the volume of such cracks, causing liquids trapped in them to depressurize rapidly and precipitate out their minerals, say the researchers.

This process, known as flash vaporization, could happen repeatedly in geologically active and gold-rich areas, such as Indonesia and South Africa.

*Nature Geosci.* <http://dx.doi.org/10.1038/ngeo1759> (2013)

For a longer story on this research, see [go.nature.com/6jwoxn](http://go.nature.com/6jwoxn)

## ZOOLOGY

## Frog feet share human hair origin

The toe pads that allow tree frogs to cling to slippery surfaces share their origin with human hair.

Toe pads are complex structures that are held in shape by a rigid network of proteins. Wim Vandebergh and his colleagues at the University of Brussels found 200 gene sequences expressed in the toe pads of the tree frog *Hyla cinerea* (pictured), but nowhere else in its toe.



ROLF NUSSBAUMER/NATUREPL.COM

The authors identified five keratins that seem to be the main structural components of toe pads; the genes that encode these five proteins have the same evolutionary origin as those for the keratins found in mammalian hair follicles.

The authors say that these proteins must have arisen in an early tetrapod ancestor before diverging to become hair in mammals and toe pads in frogs.

*Biol. Lett.* 9, 20130051 (2013)

**NATURE.COM**

For the latest research published by Nature visit:

[www.nature.com/latestresearch](http://www.nature.com/latestresearch)



# SEVEN DAYS

The news in brief

## RESEARCH

### Mars find

NASA scientists reported on 12 March that the agency's Curiosity rover had found evidence that Mars could have supported life billions of years ago. Rock samples collected by Curiosity contained clay that probably formed in salt water that was of neutral pH. See [go.nature.com/ebzeey](http://go.nature.com/ebzeey) and page 292 for more.

### Methane success

Japan announced on 12 March that it had extracted methane from methane hydrate deposits about 300 metres under the seabed. Methane, the main constituent of natural gas, has previously been obtained from deposits under frozen tundra, but marine methane hydrate reservoirs are much larger. The successful effort by the deep-sea drilling vessel *Chikyu* marks the beginning of a two-week pilot project that will assess whether methane hydrates could become an economically feasible source of energy for Japan.

## BUSINESS

### GSK absorbs Sirtris

Sirtris Pharmaceuticals, which is pursuing drugs that combat diseases related to ageing, is to be absorbed by GlaxoSmithKline (GSK), a pharmaceutical firm headquartered in London. GSK acquired Sirtris, based in Cambridge, Massachusetts, in 2008 for US\$720 million, but allowed it to function as a distinct unit. On 12 March, however, GSK confirmed that the Cambridge site would be closed and the firm's lead drug candidate — a molecule believed to activate a protein called sirtuin 1 — folded into GSK's internal drug-development programme. See [go.nature.com/nhjjtn](http://go.nature.com/nhjjtn) for more.

### Seabed mining

UK Seabed Resources in London, a subsidiary of aerospace firm Lockheed Martin UK, has been granted permission to harvest metals from the sea floor at a site in the Pacific Ocean around 1,500 kilometres southwest of Mexico, the company announced on 14 March. See page 294 for more.

### AstraZeneca cuts

Drug giant AstraZeneca says that it will cut some 1,600 staff — of which 1,300 are researchers — in a restructuring plan that will see scientists concentrated at three major research sites: Cambridge, UK (which will become the firm's new global headquarters), Gaithersburg

in Maryland and Mölndal in Sweden. The company will close down all research at Alderley Park, UK. The plan, announced on 18 March, will cost US\$1.4 billion and will see the firm spend an additional \$500 million on the UK facility in Cambridge.

### Drug-data deceit

A UK researcher has for the first time been found guilty of manipulating safety data in experiments that supported clinical trials, the United Kingdom's drug-regulatory agency announced on 12 March. Working for the drug-development firm Aptuit in Edinburgh, Steven Eaton had since 2003 selectively reported figures that were later used to assess

the concentration of a drug in the blood. The manipulation meant that failed experiments were deemed to have worked, the Medicines and Healthcare products Regulatory Agency said. He will be sentenced in April for violating regulations set out in 1999 for good laboratory practice.

## POLICY

### Energy fund

US President Barack Obama on 15 March proposed the creation of an Energy Security Trust to fund research and development into low-carbon transportation alternatives. The trust would be supplied with US\$2 billion over 10 years by diverting a fraction of the federal royalties from oil and



CROSSRAIL

## Plague graves unearthed in London

A medieval plague pit has been unearthed during construction of a Central London rail project, archaeologists announced last week. The roughly 660-year-old mass-burial site in Charterhouse Square was established to cope with the arrival in London of the Black Death — the bubonic plague. Researchers

previously obtained the genome sequence of *Yersinia pestis*, the bacterium responsible for the plague, from human remains uncovered in a separate London burial pit (see *Nature* 478, 444–446; 2011). The latest find might bring further insight into the history and genetics of the disease.

gas leases, which are poised to grow in the coming years. See page 281 for more.

## EU budget battle

The level of funding for the European Union's Horizon 2020 research programme is uncertain again after the European Parliament voted against the proposed seven-year general budget on 13 March. Heads of state agreed at a meeting on 8 February on an austerity budget of €960 billion (US\$1.2 trillion) for 2014–20, which includes about €70 billion for Horizon 2020. The parliament, which last November suggested a research budget in excess of €100 billion, approved the spending ceiling but wanted unspent funds from one area to be spent on projects in another rather than returned to member states, as well as a review of the budget in 2014.

## Species protection

Multiple species of sharks and rays, including the oceanic whitetip shark (**pictured**), were granted enhanced protection at a meeting of parties to the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) in Bangkok last week. Various species of tropical hardwood trees were also added to appendix II of CITES, which places restrictions on their



trade. Delegates accepted proposals to increase the forensic analysis of seized ivory to fight poaching of African elephants, but rejected enhanced protections for polar bears. See [go.nature.com/r076uc](http://go.nature.com/r076uc) and page 281 for more.

## Drinks ban busted

A judge struck down a ban on the sale of large servings of sugary drinks in New York city on 11 March. Enacted by city mayor Michael Bloomberg's Board of Health, the ban would have gone into effect on 12 March. But after beverage makers, grocers' groups and restaurateurs sued, Justice Milton Tingling of the State Supreme Court ruled against it, calling the ban "arbitrary and capricious". Bloomberg said his administration would appeal against the ruling. The ban would have limited restaurants, cinemas and

kiosks to selling sugary drinks in portions of 16 US fluid ounces (473 millilitres) or less.

## Mars mission boost

Russia is partnering with the European Space Agency (ESA) on its ExoMars missions to the red planet. Roscosmos, Russia's space agency, will provide the launch vehicles and scientific instruments for both a planned orbiter in 2016 and a rover in 2018, as well as the rover's descent module and surface platform. The two agencies formalized the arrangement on 14 March. NASA, which in February 2012 backed out of its original plan to partner with ESA on ExoMars, will contribute telecommunications and other support.

### PEOPLE

## Generics limbo

The head of generic drugs at the US Food and Drug Administration, Gregory Geba, resigned on 15 March after just eight months in the post. He leaves the Office of Generic Drugs just as it is being reorganized and takes on extra responsibilities. Legislation last year allowed the office to start collecting fees on applications for the approval of generic drugs, in the hope that the added resources would allow it to shorten the time from application to approval.

### FUNDING

## Spain fund freeze

The European Science Foundation (ESF) has temporarily frozen most funding for networking events planned beyond June to researchers from Spain's National Research Council (CSIC) and Ministry of Economy and Competitiveness until the agencies pay late dues. Spain has also cut payments to CERN, Europe's particle-physics laboratory near Geneva in Switzerland, and to the European Space Agency among others. The CSIC and ESF say that a resolution is forthcoming.

### CORRECTION

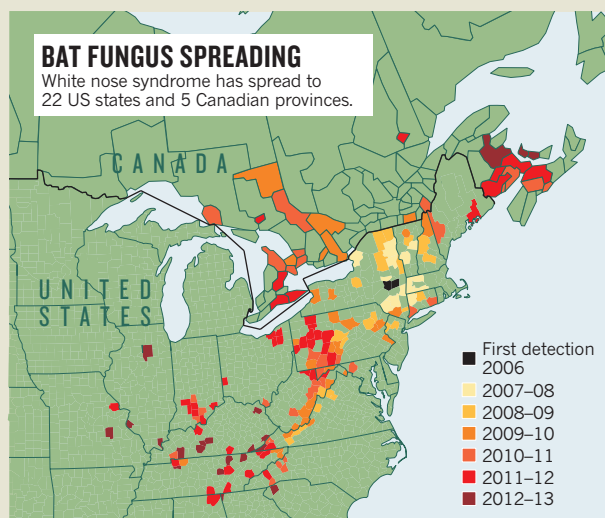
The story 'Carbon spike' (*Nature* **495**, 146; 2013) wrongly stated that the atmospheric concentration of CO<sub>2</sub> has risen by 70% from pre-industrial levels. The global average of 395 p.p.m. represents a 41% increase.

### NATURE.COM

For daily news updates see:  
[www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

White nose syndrome, a fungal disease of bats that is sweeping through North America, was this month documented in two new states: Georgia and South Carolina. The deadly infection, named for the powdery fungal growth on bats' muzzles, probably came from Europe, where the fungus (*Geomyces destructans*) is endemic. Since it first appeared in New York in 2006, the fungus has spread to 22 states and 5 Canadian provinces, and has killed an estimated 7 million bats.





# NEWS IN FOCUS

**ENVIRONMENT** Shale-oil boom carries extra carbon cost **p.290**

**SPACE** Mars-rover team feels pressure to pick up the pace **p.292**

**RESOURCES** Will deep-sea mining finally take the plunge? **p.294**

**RESEARCH** Marathon experiments, and their lasting rewards **p.300**



JANET MCCONNAUGHEY/AP



Chimp Haven in Louisiana is the final destination for retired NIH research animals.

## ETHICS

# Time called on chimp work

*NIH likely to retire hundreds of government-owned chimpanzees.*

BY MEREDITH WADMAN

**G**overnment-funded chimpanzee research in the United States has been on notice for more than a year. This month, its time is probably up.

First came a report from the Institute of Medicine (IOM) in December 2011, which found that most chimp research is unnecessary, given that almost all questions can be explored in other animal models or *in vitro*.

Now Francis Collins, director of the National Institutes of Health (NIH), is expected to accept the bulk of the recommendations from an NIH working group that would retire several hundred agency-owned chimpanzees, leaving a core of just 50 animals available for

NIH research. The working group, which on 22 January released proposals for how the NIH should implement the IOM report, also said that the agency should end six of the nine invasive studies that it funds. And the other three would be allowed to continue only if the animals' environments are substantially improved, probably at great expense.

Should Collins accept the recommendations, it would be "a milestone in animal-research practice and policy, and an endorsement of the conclusion that there is a rapidly decreasing need for the use of chimpanzees in research", says Jeffrey Kahn, a bioethicist at Johns Hopkins

University in Baltimore, Maryland, who led the 2011 IOM committee.

The working group did not identify which experiments it thought should end, so NIH researchers will learn the fate of their studies only after Collins's decision. Several chimpanzee researchers contacted by *Nature* declined to speak on the record about their work, but some lamented that unique science would disappear.

The nine invasive biomedical experiments include one that is searching for an infectious trigger for primary biliary cirrhosis, a disease that slowly destroys the liver and is fatal without a transplant. Another, conducted by a biotechnology company, tests an agent that binds to molecular targets on immune cells and is intended to shut down the immune

**NATURE.COM**  
For more on the controversy over chimp research:  
[go.nature.com/ugfawq](http://go.nature.com/ugfawq)



► response that causes diseases such as rheumatoid arthritis and lupus. The candidate drug's targets occur in only two species: humans — in whom safety tests of the compound would be ethically impermissible — and chimpanzees. If the study is killed, “this line of investigation will probably have to stop”, says one observer with knowledge of the study.

Although the IOM committee was unanimous in declaring chimpanzees unnecessary for almost all invasive research, it was split on whether the animal is essential for developing a vaccine against hepatitis C, which kills some 350,000 people globally each year. The chimpanzee is the only animal model that can test whether an immune response to a candidate vaccine would protect against hepatitis C infection. One of the nine studies now in jeopardy explores why the immune response in chimpanzees often cannot clear hepatitis C infections.

“Everybody working with chimpanzees in the hepatitis C field realizes that they are really unique animals,” says Stanley Lemon, who studies hepatitis C at the University of North Carolina, Chapel Hill, and who resigned from the NIH working group last year. “But you have to balance that against the weight of human morbidity and mortality.”

The working group also recommended that five of 13 behavioural and comparative-genomics studies should end. The Yerkes National Primate Research Center in Atlanta, Georgia, hosts some of these studies, which include an assessment of the roles of the hormones vasopressin and oxytocin in social cognition, and a look at how genes and early social experiences affect brain development and behaviour.

Even if Collins allows the Yerkes studies to continue, the recommendation for new housing requirements could doom the projects. The working group proposed the provision of 6.1 vertical metres of climbing structure and 93 square metres of space for each animal; fresh materials for daily construction of nests; foraging opportunities; and housing in social groups of at least seven. “We estimate it could cost tens of millions of dollars for our centre alone to renovate space that will have limited use given our ageing colony,” says Yerkes spokeswoman Lisa Newbern.

Some ethicists say that if Collins embraces the recommendations, his decision will not only mark the end of most government-supported chimpanzee research, but also raise questions about work on other non-human primates. The NIH supports a network of primate research centres that houses some 27,000 non-human primates of 20 species. “We’re not going to suddenly decide that all primates have got to be retired,” says Gregory Kaebnick, a research scholar at the Hastings Center in Garrison, New York, who last month led a study on changing ethical views on animal research. “But we might find ourselves trying to reassess the way we’ve been using them.” ■



Natural gas from oil wells in North Dakota is often burned rather than captured and sold.

#### ENERGY

# Oil boom raises burning issues

*Unburned methane could be adding to the environmental impact of gas flares in North Dakota.*

BY JEFF TOLLEFSON

When Paul Shepson flew his twin-propeller Beech aircraft over the Williston Basin in North Dakota last June, the perverse economics of a modern-day energy bonanza were on vivid display. Shepson, an atmospheric chemist at Purdue University in West Lafayette, Indiana, saw 10-metre-high flames writhing in the wind as shale-oil producers burned off the natural gas that was coming up along with the oil — a valuable resource that here, far from gas pipelines, is just a nuisance.

In 2011, North Dakota's oil producers flared around 1.4 billion cubic metres of natural gas, according to the Energy Information Administration (EIA) in Washington DC. That was nearly 32% of the total amount of gas produced by the state, and enough to heat around 700,000 houses. And the flaring increased in 2012. Scientists are just beginning to assess how the carbon dioxide from the flames adds to the climate footprint of the oil, and how black carbon and other pollutants affect local air quality.

But Shepson has also begun to detect evidence of a hidden threat: plumes of pure methane — a much more potent greenhouse gas than CO<sub>2</sub> — that is either leaking from the oil and gas infrastructure or being vented by extractors who are not flaring. Although he has yet to quantify the methane emissions, Shepson says that the energy coming out of the basin could have a steeper environmental price than many realize.

Over the past several years, companies have been using the same hydrofracturing technologies that have spurred a broader US gas rush to drill thousands of oil wells into the Bakken shale formation, which stretches across North Dakota, Montana and the Canadian province of Saskatchewan. High oil prices are driving the boom, which some predict could help to make the United States the world's largest oil producer by 2020. Companies that build pipelines and gas-processing plants simply cannot keep up.

Some 65% of the oil from the Bakken shale is now being shipped out by rail. But the raw natural gas that bubbles up along with it — a mixture of lighter liquid hydrocarbons and

DANIEL ACKER/BLOOMBERG VIA GETTY



gases such as propane, butane and methane—requires both pipelines and processing, says Rusty Brazier of RBN Energy, a consultancy based in Houston, Texas. Some of the liquid gas is shipped to Canada, where it is used to dilute raw bitumen from the Alberta tar sands for shipment back to the United States, but production has so far outpaced demand. “The companies would love to sell that natural gas,” says Rich Alonso, a partner at Bracewell & Giuliani, an industry lobbying firm based in Washington DC. But “it has got no place to go”.

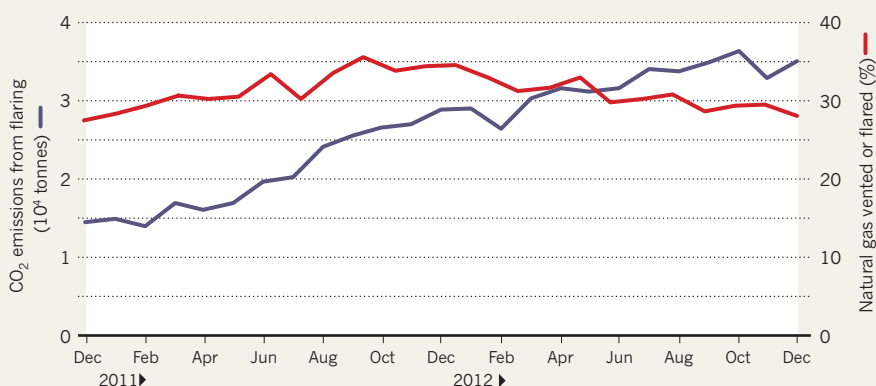
### BURNING PROBLEM

The upshot is flaring, and lots of it. North Dakota’s flaring rose more than sixfold between 2006 and 2011, and total US flaring increased by 62% over the same period, according to the EIA. Satellite data from researchers at the National Oceanic and Atmospheric Administration in Boulder, Colorado, show that even before the shale-oil boom, the United States was flaring more gas than all but a handful of countries, including Russia, Nigeria, Iran and Iraq.

Environmentalists are pushing North Dakota to halt flaring and require that the gas be used to produce electricity or be processed into usable products, such as fertilizer. Scientists and entrepreneurs are working on mobile

### UP IN SMOKE

The proportion of the natural gas being vented or flared during shale-oil operations in North Dakota hovers at around one-third, but carbon dioxide emissions have grown with booming production.



SOURCE: CLEAN AIR TASK FORCE/EIA

refineries that could be deployed in the field to process the stranded gas. And at the federal level, environmentalists are asking the US Environmental Protection Agency to regulate methane emissions under the Clean Air Act. But there is no obvious resolution in sight.

“It’s being treated as an infrastructure problem, but I see it first and foremost as a regulatory problem,” says Brad Crabtree, policy director for the Great Plains Institute, an environmental advocacy organization based in Minneapolis, Minnesota. “The industry can afford to manage the gas and should be required to do so.” Industry officials say that gas infrastructure is already expanding and argue that regulations would only slow that development.

Shepson initially focused his investigations on the flares, hoping to measure the pollutants they emit and how much methane was escaping unburned. He and his crew flew through the plumes rising from 11 major flares, and found that they were highly efficient, burning all but around 0.1% of the methane. But the team also documented numerous plumes of pure methane, suggesting that natural gas is escaping directly into the atmosphere in many locations. This is particularly troubling because, molecule for molecule, methane has roughly 25 times the warming power of CO<sub>2</sub>.

Exactly how much methane is escaping remains unclear, however, as does its source, although some suspect that flaring is not as universal a practice as believed. “We don’t have an estimate for how leaky the Bakken is in terms of methane emissions, but if you are going to flare

the gas anyway you probably aren’t going to be too careful,” says David McCabe, an atmospheric scientist at the Clean Air Task Force, an environmental group based in Boston, Massachusetts, which funded the work.

Scientists are just beginning to incorporate this information into life-cycle greenhouse-gas assessments of shale-oil developments. The oil coming out of the Bakken formation is high-quality light crude, which produces fewer CO<sub>2</sub> emissions than thick oil from the Canadian tar sands or Venezuela. But no one knows for sure how natural-gas leakage and flaring affect that

**“If you are going to flare the gas anyway, you probably aren’t going to be too careful.”**

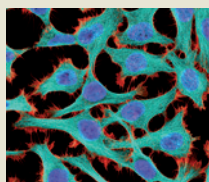
equation. Assuming a best-case scenario in which all the methane is combusted, McCabe estimates that the flares would have generated nearly 4 million tonnes in

CO<sub>2</sub> emissions over the course of 2012 (see ‘Up in smoke’), akin to running a medium-sized coal-fired power plant in the prairie. The leaking methane could add substantially to that climate impact.

Shepson says he hopes to get a clearer reading of the methane leakage during a second round of investigations this year, but he is worried about the big picture. “When you fly over this region, what hits you is the scale of the development,” he says. “We are reinvesting in an aggressive way in fossil fuels, and the price of all of that will be a more certain and larger impact from climate change.” ■ SEE EDITORIAL P.281



### TOP STORY

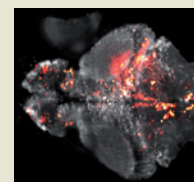


Sequencing finds HeLa cell genome riddled with errors  
[go.nature.com/inxzuw](http://go.nature.com/inxzuw)

### MORE NEWS

- Super-dense extrasolar objects could be new type of planet [go.nature.com/3tti6q](http://go.nature.com/3tti6q)
- Early birds flew on four wings [go.nature.com/Gzem1p](http://go.nature.com/Gzem1p)
- Obama unveils US\$2-billion plan to get US cars off petrol [go.nature.com/zdzpyd](http://go.nature.com/zdzpyd)

### VIDEO



Whole-brain imaging shows fish's activity down to single neurons  
[go.nature.com/viqdja](http://go.nature.com/viqdja)

MISHA B. AHRENS & PHILIPP J. KELLER



# Mars rover under pressure to reach mountain goal

*Curiosity's memory glitch prompts mission scientists to pick up the pace.*

BY ALEXANDRA WITZE

Seven months after it hurtled to a landing on Mars, NASA's Curiosity rover has rolled just 738 metres across the surface. It remains about 10 kilometres from its primary goal: Aeolis Mons, a 5-kilometre-high mountain thought to contain layers formed in ancient ponds. With the mission nearly one-third of the way through its allotted two years, the science team is beginning to feel the pressure to get moving.

If Curiosity were to die tomorrow, “we could certainly say we achieved a lot”, says science-team member Laurie Leshin, a planetary geochemist at Rensselaer Polytechnic Institute in Troy, New York. “But none of us would be satisfied.”

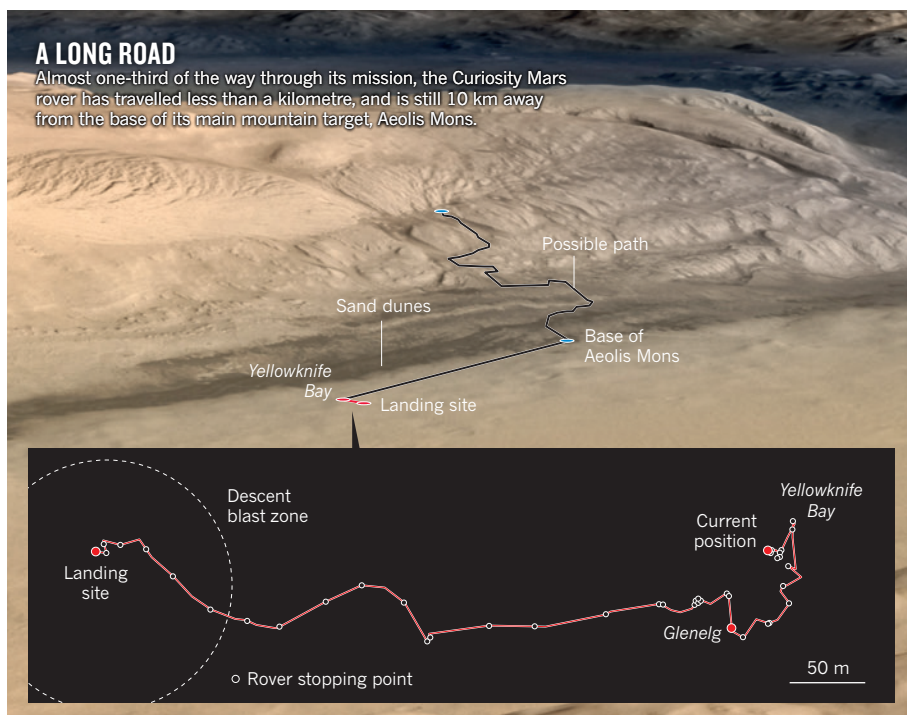
It took two months just to unfurl the array of scientific instruments on the US\$2.5-billion Curiosity — the biggest, most expensive and most lavishly equipped rover ever sent to Mars. It is also planetary exploration's most complicated robot ever, with about 400 scientists working on ten different instruments: a social environment that can slow things down. But the first serious glitch of the mission — a memory malfunction on 28 February that forced operators to use the rover's back-up computer — has reminded Curiosity's handlers that the machine is mortal.

From the start, however, mission scientists have found plenty of intriguing distractions. Soon after touching down in Gale Crater on 6 August last year, the rover found evidence of a dried-up stream bed near the base of Peace Vallis, an 80-square-kilometre fan of sediments that washed down from the walls of the crater.

In October, after heading to Glenelg, a junction where three terrains come together, the rover stopped for almost two months at a drift of sand dubbed Rocknest. Here, Curiosity plunged its sampling scoop into the sand and tipped the material into its two main analytical instruments, CheMin (for Chemistry and Mineralogy) and SAM (Sample Analysis at Mars). On 19 March, an entire session of the Lunar and Planetary Science Conference at the Woodlands, near Houston, Texas, was devoted to the findings at Rocknest. The dune turns

► NATURE.COM

For more on the Mars Curiosity rover, see *Nature's special:* [nature.com/curiosity](http://nature.com/curiosity)



SOURCE: NASA/JPL-CALTECH/ESA/DLR/FU BERLIN/MSS

out to be chemically similar to soils analysed elsewhere on Mars, suggesting that materials such as water move between the soil and atmosphere in similar ways around the planet.

Also at the conference, scientist Chris Webster of NASA's Jet Propulsion Laboratory in Pasadena, California, reported the latest on methane measurements taken by the rover's atmosphere-sniffing, tunable laser spectrometer. The spectrometer team's findings have lowered the upper limit for methane in Mars's atmosphere — to as little as 3 parts per billion by volume. The scarcity of methane casts doubt on the possibility that Martian microbes are emitting the gas into the atmosphere (see *Nature* **491**, 174; 2012). The instrument has been used just three times in the first 200 sols (Martian days), even though it can be activated at night when there is less demand from other science instruments on the rover. This is a sign of the intense competition for rover time: one of the reasons for the slow pace.

Since 23 January, the rover has more or less stayed in one place, snuggled up next to a reddish rock nicknamed John Klein, in a region called Yellowknife Bay. At John Klein, Curiosity

used its drill for the first time, and analysis of the powdered drilling sample revealed clay minerals that formed in a watery environment, along with oxidized and reduced forms of sulphur. That suggests that Yellowknife Bay was once a benign environment for microbial life (see *Nature* <http://doi.org/ktt>; 2013). Yellowknife “has been a great place to start”, says team member Roger Wiens of Los Alamos National Laboratory in New Mexico. “We wouldn't have wanted to miss this for anything.”

But Jorge Vago, project scientist for the European Space Agency's ExoMars mission, says that he is already thinking about how he will need to do things faster with his rover, which is scheduled for launch in 2018. “Clearly, we couldn't afford to run ExoMars in the way [Curiosity] is doing the mission,” he says. Instead, ExoMars is likely to have a shorter mission length, so might need to perform its science at a pace somewhere between the tortoise speed of Curiosity and the hare-like rate of NASA's Spirit and Opportunity rovers, which landed on Mars in 2004 with primary missions of a little over three months.

Just as Curiosity's science team is getting



itchy to move the rover again, however, it will have to take another long break. For most of April, Mars will be behind the Sun as seen from Earth, and no spacecraft on or around Mars will be able to radio home. After Mars emerges, the team aims to drill into the John

Klein rock once more, to repeat the measurements. Then the rover will finally hit the road to Aeolis Mons, also known as Mount Sharp.

Curiosity will follow a seemingly flat, unobstructed path around a set of sand dunes (see 'A long road') at speeds of up to 100 metres a day,

says Raymond Arvidson, a planetary geologist at Washington University in St Louis, Missouri. "I call it the rapid transit route," says Arvidson. "There's general agreement that we have to get to the base of Mount Sharp before the primary mission ends, with a lot of time to spare." ■

## GENOMICS

# Gene-analysis firms reach for the cloud

*Online bioinformatics companies rush to provide genomics platforms and software for hospitals.*

BY ERIKA CHECK HAYDEN

For Chaim Jals at the Center for Rare Jewish Genetic Disorders in New York, DNA sequencing is the easy part. It costs less than US\$1,500 per person to have the important parts of his clients' genomes sequenced. But it would be dauntingly expensive to maintain servers and staff to analyse the data and identify mutations that might be causing the undiagnosed diseases that afflict his clients' families.

So Jals, the centre's director of genetics resources and services, has outsourced parts of the analysis. He uploads his clients' sequencing data to cloud-computing software platforms, where he can run analyses without having to set up the infrastructure in-house. The cost is about \$100 per person. And the cloud-based interfaces let him collaborate with doctors in Israel without worrying about repeatedly transferring data on slow Internet connections. "For me, it is convenient and cost-efficient," he says.

Jals and the way he works represent a new and mostly untapped market for a new crop of genetics interpretation and analysis firms, which will be touting for customers at a meeting of the American College of Medical Genetics and Genomics in Phoenix, Arizona, on 19–23 March. Dozens of these firms have emerged, some in the past year, as ever more affordable sequencing moves from academia into the clinic (see *Nature* 494, 290–291; 2013). Doctors will increasingly want to use sequencing data to guide decisions about patient care, but might not necessarily want to invest in staff and software to make sense of those data.

"It's a huge unmet need," says David Ferreiro, a biotechnology analyst with investment bank Oppenheimer & Company in New York, which invests in many sequencing-technology and -analysis companies.

Where there is a need, there is also money to be made. The companies, many of them based

in California's Silicon Valley, have been tempted by a market in outsourced sequencing and analysis software that by 2016 could top \$4 billion per year, according to BCC Research, a market-research company in Wellesley, Massachusetts (See 'Genes in the marketplace'). "The sky's the limit," says Andreas Sundquist, chief executive of DNAnexus in Mountain View, California, which provides genetic-analysis software on its cloud-based platform and allows users to upload and run their own algorithms.

Other firms offer a range of approaches. Seven Bridges Genomics, based in Cambridge, Massachusetts, aims to be accessible to people with no expertise in bioinformatics, and provides access to free tools for designing custom-made analysis pipelines. Ingenuity Systems in Redwood City, California, allows users to upload a list of mutations in a person's genome, and finds those most likely to cause disease. Personalis, down the road in Menlo Park, offers sequencing services and interpretation for clinicians and pharmaceutical and biotechnology companies. Last week, the company won a \$1.53-million contract with the US Department of Veterans Affairs to look for genetic variants

in samples from as many as one million military veterans, to explore the variants' roles in disease. (The company will outsource the sequencing to Illumina, a market leader in sequencing technology based in San Diego, California.)

The activity is reminiscent of that a decade ago, when bioinformaticians started up a flurry of companies, most of which were unsuccessful because the path from a genetic-disease marker to a profitable drug has not been straightforward. Today's companies have moved on to other challenges, says Steven Brenner, a computational geneticist at the University of California, Berkeley. He says that they will have to prove that their products are better than freely available software — and do so without disclosing their intellectual property. "These companies have a tricky situation," he says.

And they are facing competition from the sequencing-technology firms themselves. Last year, Illumina opened BaseSpace Apps, a marketplace for online analysis tools to be used on data uploaded to the company's own cloud-computing platform. Information-technology companies are getting in on the action, too: Oracle, based in Redwood Shores, California, has its own products aimed at helping researchers and hospitals to analyse data.

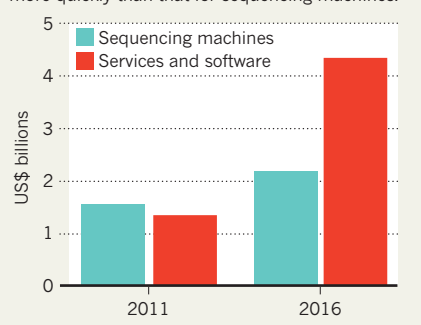
But one of the biggest questions will be how deeply analysis companies can reach into medical settings, where privacy concerns are paramount. Hospitals can be fined if patient privacy is compromised, and clinical geneticists may be uneasy about uploading data to the cloud. "It's your licence and your lab that go on the line when it comes to reporting a clinical result," says Elizabeth Worthey, director of genomic informatics at the Human and Molecular Genetics Center at the Medical College of Wisconsin in Milwaukee. That is a large part of why many hospitals have so far chosen to build their own analysis infrastructure, says Ferreiro.

To address privacy concerns, Bina Technologies in Redwood City sells a server that can sit in a customer's own data centre and is optimized to run genome-analysis software. Knome of Cambridge, Massachusetts, announced last year that it plans to sell \$125,000 genome-analysis machines for use in customers' labs (see *Nature* 490, 157; 2012).

It seems unlikely that any single analysis company will rule the market; the range of customers who need to interpret sequence data is growing, and each has their own needs. "We would rather see a thousand flowers bloom," says Sundquist. "I don't think we're going to see consolidation anytime soon." ■

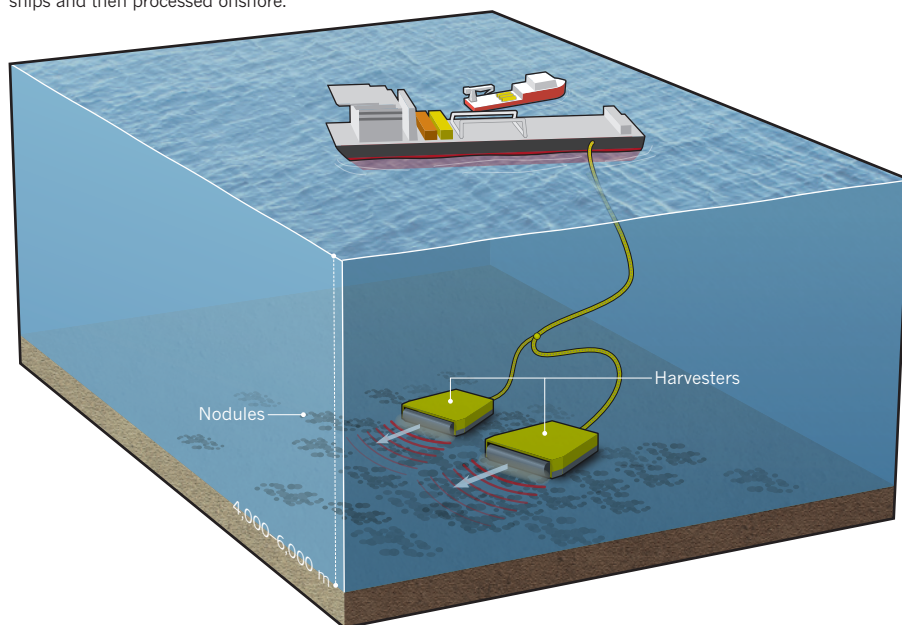
## GENES IN THE MARKETPLACE

The market for outsourced gene-sequencing work and custom-made software is projected to grow more quickly than that for sequencing machines.



## WEALTH FROM THE SEA

Systems are being developed to harvest mineral-rich nodules from the sea floor. In this scheme, the harvester would vacuum or rake up the nodules, grind them and mix them with sea water. The slurry would be piped to ships and then processed onshore.



## MINING

# UK company pursues deep-sea bonanza

*After decades of preparation, firm moves forward with plans to harvest mineral-rich nuggets from sea floor.*

BY MARK SCHROPE

A bizarre episode from the cold war could foreshadow a new era of seabed mining. In 1974, the US government used a ship owned by eccentric billionaire Howard Hughes to pick up the remains of a Russian nuclear-missile-bearing submarine that sank in more than 5,000 metres of water. As cover, the United States claimed that the expedition was collecting polymetallic nodules: rocky lumps that litter the deep sea floor, rich in metals such as manganese, nickel, copper, cobalt and rare-earth elements.

It turns out that Lockheed, the US defence contractor that made the equipment used to grab the sub, had a real interest in the minerals. And last week, its London-based affiliate Lockheed Martin UK announced that it is to explore the harvest of nodules in the Pacific Ocean, 1,500 kilometres southwest of Mexico.

If it were to go ahead, the mining would be the culmination of decades of technological effort and exploration. “Lockheed Martin (US) in the end did more development and at-sea testing than any other company,” says Caitlyn

Antrim, executive director of the Rule of Law Committee for the Oceans in Washington DC, and a deep-seabed-mining consultant. The venture could add £40 billion (US\$60 billion) to the UK economy over the next 30 years, said UK Prime Minister David Cameron last week.

Cameron’s government sponsored Lockheed’s claim to the 58,000-square-kilometre potential mining site through the company’s subsidiary UK Seabed Resources in London. The International Seabed Authority in Kingston, Jamaica, which granted the exploration rights, has also granted claims in the region to several other countries, but work towards harvesting has been slow. The UK effort has advantages, says Antrim: Lockheed has proven technologies and the most nodule-bed data.

Polymetallic nodules form over thousands of years on the sea floor, through processes that are still not fully understood; most nodules range in size from the diameter of a golf ball to that of a large potato. Between the 1960s and the 1980s, various companies explored the possibility of harvesting them — but as metal prices fell, the prospect became economically unfeasible. In recent years, however, climbing

prices and advances in deep-sea equipment such as remotely operated vehicles and flexible risers — pipes used to lift material — have revived interest. “It’s part of the art of the possible now,” says Miles Libbey, a former employee of Lockheed Martin US, and the founder of Seaminr, a seabed-mining consultancy in Urbana, Virginia.

Harvesting nodules is technically easier than mining the mineral-rich remnants of non-active hydrothermal vents on the sea floor, a proposal that has also garnered attention (see *Nature* 447, 246–247; 2007). Nodules rest unattached, and often visible, in sea-floor sediments, so obtaining them would not entail grinding up rocky sea-floor crust. Lockheed has not released specific plans, but the general scheme for harvesting uses a device that moves along the sea floor, vacuuming or raking up nodules. The equipment would grind down the nodules to create a slurry that would be pumped to ships and then transported onshore for processing (see ‘Wealth from the sea’).

The most lucrative product would probably be nickel, which makes up only about 1% of the nodules but is used in everything from stainless steel to rechargeable batteries. Rare-earth elements are much less concentrated at around 0.1%, but will still be valuable: they are crucial for technologies in the aerospace and alternative-energy industries. China currently controls the vast majority of supplies.

Philip Rood, a spokesman for Lockheed Martin UK, says that it hopes operations can begin in 5–6 years, once Lockheed has completed technology tests and environmental studies.

Data are so far sparse on the degree to which the operations would threaten deep-sea life such as sediment-dwelling sea cucumbers, worms and small crustaceans, or creatures that live higher in the water column, such as fish. A study at a nodule plain off Peru found that sea-floor communities showed limited recovery after disturbance, with sediment dwellers gradually returning over 7 years (H. Bluhm *Deep-Sea Res. II* 48, 3841–3868; 2001).

Craig Smith, a deep-sea biologist at the University of Hawaii at Manoa, will lead an initial assessment of sea-floor life for Lockheed’s project, gathering baseline data for the potential harvest zone and comparison areas. “I think this is going to happen,” he says of nodule harvesting. “To manage the environmental impacts well we really need excellent information.”

Cindy Van Dover, director of the Duke University Marine Laboratory in Beaufort, North Carolina, has worked on assessments of deep-sea mining impacts. She says, “I don’t think there are red flags but I do think there are yellow flags.”

Her main concern is one of scale. “One tiny patch is not going to make any difference but what I can’t figure in my head is what happens when that’s multiplied.” ■

PAUL JACKMAN/NATURE

## PUBLISHING

# Text-mining spat heats up

*Scientists and publishers clash over licences that would let machines read research papers.*

BY RICHARD VAN NOORDEN

It is seen as the future of computer-based research — if only the gatekeepers would let scientists in. Researchers have complained bitterly over the past year that publishers do not allow them to use computer programs to download and crawl across the text of research articles, a methodology known as text mining that can reveal large-scale patterns in the studies (see ‘Uses of text mining’).

Fearful that their content might be freely redistributed, publishers tend to block programs that they find crawling the full text of articles, making no exceptions for users who have paid for access. They give permission only on a case-by-case basis to those who negotiate agreements on access and use. Now, the European Commission (EC) and publishers’ consortia are trying to craft clearer rules. But complaints last month to an EC group set up to discuss text and data mining suggest that disagreements are still rife.

“Data- and text-mining techniques... could hold the key to the next medical breakthrough, if only we freed them from their current legal tangle,” Neelie Kroes, vice-president of the European Commission, told a Brussels intellectual-property summit last September.

Publishers say that so far, few researchers are asking permission to mine text. Still, Amsterdam-based publisher Elsevier says that computer robots crawling its ScienceDirect site made up 4% of total web traffic on the platform in 2012, almost twice the level in 2011. Whatever the intention of such programs, the figure suggests that machines, not just humans, are increasingly poring over articles.

Raul Rodriguez-Esteban, a computational biologist at drug company Boehringer Ingelheim in Ridgefield, Connecticut, says that he ran 160 text-mining queries in 2012. In one, he searched more than 23,000 articles to pick out hundreds of proteins that could relieve a mouse model of multiple sclerosis. He then sketched a network of other proteins that interacted with them, and found new potential drug targets. Academic researchers covet this capability, but say that it takes months or years to negotiate agreements. It took Max Haeussler at the University of Santa Cruz, California, three years to get the rights to download 3 million articles, from which he extracts DNA data to annotate an online map of the human genome (see *Nature* **483**, 134–135; 2012).

Later this year, the United Kingdom will make text mining for non-commercial purposes exempt from copyright, potentially

## A RICH RESOURCE

### Uses of text mining

#### Linking genes to research papers.

The text2genome project pulls out DNA sequences from around 3 million research papers to produce an online genome map in which each region is linked to relevant articles.

[go.nature.com/iupijx](http://go.nature.com/iupijx)

**Mapping the brain.** The NeuroSynth project extracted brain-scan data from almost 4,400 research articles, allowing users to link locations in the human brain with associated research terms and topics. [neurosynth.org](http://neurosynth.org)

**Chemistry data.** SureChem (owned by the same parent company as *Nature*) harvests and makes freely available data on molecules from some 20 million patents. [surechem.com](http://surechem.com)

**Drug discovery.** Researchers searched free abstracts from more than 20 million articles in the MEDLINE database, and discovered an indirect link between E-cadherin (a cell-adhesion molecule) and Parkinson’s disease. [go.nature.com/fsv4az](http://go.nature.com/fsv4az)

allowing scientists to mine any content they have paid for. Text miners want the EC to do the same. “The slogan that is doing the rounds among users currently is: ‘the right to read is the right to mine,’” says John McNaught, deputy director of the National Centre for Text Mining at the University of Manchester, UK.

But the EC’s working group to discuss text and data mining, set up this year, has already run into controversy. After a meeting on 4 February, researchers and librarians complained that the group was discussing only how to work with text-mining licences, not how to exempt text mining from copyright. “This will only raise barriers to the adoption of this technology and make computer-based research in many instances impossible,” they wrote on 26 February in a letter to Kroes and three other European commissioners, who have not yet responded.

A second meeting on 8 March provided little assurance that exceptions would be seriously considered, says Ross Mounce of the University of Bath, UK, who is using text mining to

extract trees of evolutionary relationships from the literature. The EC working group hopes to reach a conclusion by the end of the year.

The situation is better in the United States, where some lawyers think that text mining might be permitted by ‘fair use’ rights, which allow snippets of text to be freely copied. But no one knows for sure, and many researchers are wary of testing the bounds of this legal grey area.

Some publishers say that unrestricted text mining could strain their servers, and so agreements will always be needed to specify when and how articles may be downloaded. CrossRef, a non-profit collaboration of thousands of scholarly publishers, is developing a system to let researchers agree to standard text-mining terms by clicking a button on a publisher’s website. CrossRef’s Geoff Bilder hopes that the system will roll out by the end of the year.

The Copyright Clearance Center (CCC) in Danvers, Massachusetts, which works with publishers on rights licensing, is pursuing a more ambitious effort. It would act as an intermediary, collecting publishers’ terms and content and storing them on a website for researchers, says the CCC’s Roy Kaufman. It is working with six publishers (including Nature Publishing Group) and with drug and chemical firms eager to mine the literature.

Heather Piwowar of the National Evolutionary Synthesis Center in Durham, North Carolina, who studies how researchers use data, says that it is unfair that large firms such as Google are allowed to crawl across content to index it — yet scientists are restricted. “Is this defensible on the grounds that Google knows what it is doing but The Rest Of Us Can Not Be Trusted?,” she blogs. “I sure hope not.” ■

## CORRECTION

The News story ‘DNA tool kit goes live online’ (*Nature* **495**, 150–151; 2013) wrongly located Randy Rettberg at the Massachusetts Institute of Technology. He is at the non-profit organization the iGEM Foundation in Cambridge, Massachusetts. The News story ‘Sticky problem snares wonder material’ (*Nature* **495**, 152–153; 2013) said that the 2010 Nobel prize was awarded for graphene’s discovery, but it was for experiments involving graphene. And the World View ‘The unlikely wisdom of Chairman Mao’ (*Nature* **495**, 143; 2013) wrongly stated that Mao Zedong was China’s premier.



# STAR TRACKER

As an early adopter of astronomical technology, Andrea Ghez is revealing secrets about the giant black hole at the Galaxy's centre.

BY ANN FINKBEINER

**T**he technology was a complete joy, says Andrea Ghez, thinking back to the mid-1980s and her first time helping out at an observatory. She wanted to learn everything. "How to open the dome! How to fill the instrument with liquid nitrogen! Develop the plates! Reduce the data! Coding!"

And then there was the science. Ghez did not know much at the start; she was majoring in physics at the Massachusetts Institute of Technology (MIT) in Cambridge, working for an astronomer as her undergraduate research experience. But as she learned more about his research into unusual cosmic sources of X-rays, Ghez became enthralled by the thought that some of those sources might be black holes — singular points with a gravitational pull so strong that not even light can escape them. "It got me completely fascinated by black holes," she says. By the time she had spent two undergraduate summers working at telescopes in Arizona and Chile, Ghez was hooked. "I fell in love with the whole profession."

Now an astronomer at the University of California, Los Angeles, she still feels the same. Her fascination with black holes has led her into a pioneering, decades-long study that has proved the existence of the biggest black hole in our cosmic neighbourhood: the 4.1-million-solar-mass behemoth that lies at the centre of the Milky Way<sup>1,2</sup> (see 'The monster in the middle'). This work earned her a MacArthur 'genius' award in 2008, and half of the Crafoord prize, astronomy's Nobel, in 2012.

Ghez's love of technology helps to explain why her quest has been so fruitful. Most astronomers use only the tools they know, but Ghez is an enthusiastic early adopter — first in line to try out cutting-edge detectors and optical techniques that are barely out of the laboratory. "I like the risk of a new technology," she says. Maybe it won't work. But maybe it will open a fresh window on the Universe, answering "questions you didn't even know to ask", she says. "Any time you look, you're astounded!"

Reinhard Genzel, a director of the Max Planck Institute for

Extraterrestrial Physics in Garching, Germany — the co-winner of the 2012 Crafoord prize and Ghez's sharpest competitor on the Galactic Centre work — puts it very simply. "Andrea," he says, "is one of a rare adventurous class."

## DEEP FOCUS

Ghez's devotion to her work would make her seem fierce — if she weren't always smiling, and her sentences didn't keep exploding into verbal capitals. As it is, with her barely controlled curls, straight-across eyebrows and direct gaze, she conveys a cheerful intensity. She doesn't digress when she talks; she focuses. And she has always had a certain determination.

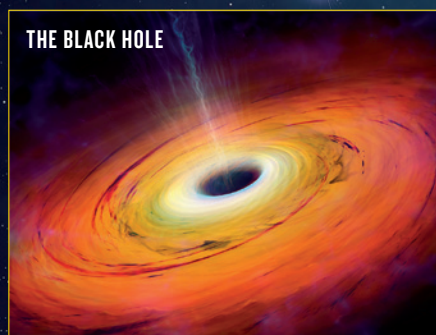
According to Ghez family legend, when 4-year-old Andrea watched the first Apollo Moon landing with her parents in Chicago, Illinois, on 20 July 1969, she announced that she, too, was going to the Moon as an astronaut. True, she also wanted to be a ballerina. But while attending the progressive University of Chicago Laboratory Schools, she says, she became "really clear" that she loved mathematics and science. That passion took her to MIT in 1983 and then, after her epiphany in the observatory domes, to the California Institute of Technology (Caltech) in Pasadena for graduate studies in astronomy.

Caltech, Ghez explains, "had the best toys by far". Among them was the 5-metre Hale Telescope, then one of the world's largest, on California's Palomar Mountain. But the toy that particularly captured Ghez's interest was an experimental speckle imager, an instrument intended to get around astronomers' eternal problem with air. Earth's atmosphere is transparent but turbulent — a collection of bubbling 'cells' that are warmer here, cooler there, and constantly moving. Looking at the sky through all that is like looking at pebbles on the bottom of a rippling stream: the light coming into the telescope

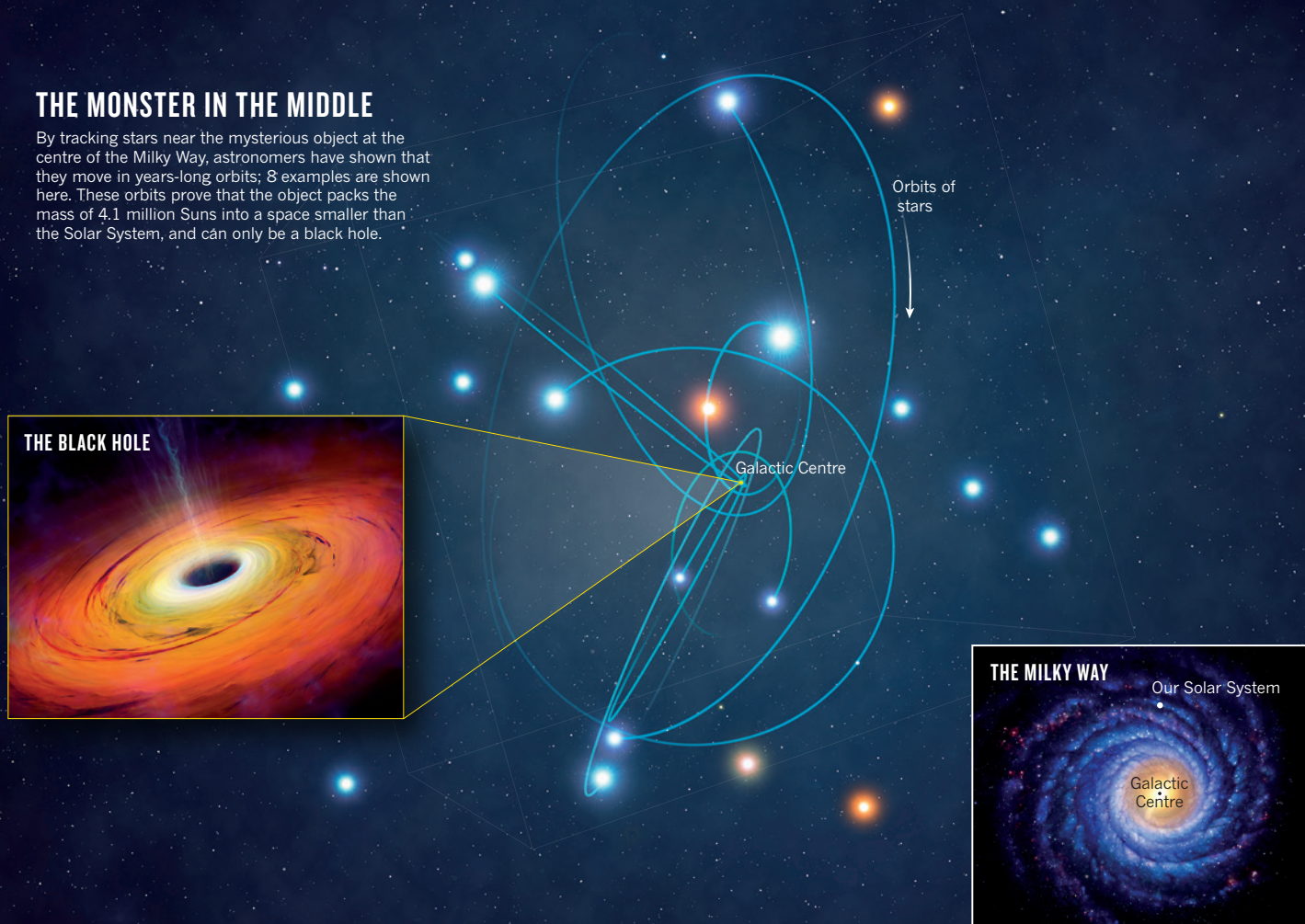
**NATURE.COM**  
Read more about the  
Milky Way at:  
[go.nature.com/uoot1n](http://go.nature.com/uoot1n)

## THE MONSTER IN THE MIDDLE

By tracking stars near the mysterious object at the centre of the Milky Way, astronomers have shown that they move in years-long orbits; 8 examples are shown here. These orbits prove that the object packs the mass of 4.1 million Suns into a space smaller than the Solar System, and can only be a black hole.



THE BLACK HOLE



THE MILKY WAY

MARK GARLICK

flickers, dances and fragments, smearing the point-like image of each star into a fuzzy ball.

Speckle imaging freezes the dancing images in place with a camera that captures very short exposures every few milliseconds, taking maybe 10,000 or more shots in total. The result is a sequence of very faint images in which the distorted light from each star produces a scattering of spots: the speckles. Computer processing recombines the speckles into one spot per star. Then all the exposures can be aligned and stacked to produce a final image with the worst of the atmospheric smearing removed.

At the time, speckle imaging was well established but done mostly at optical wavelengths and used infrequently, because it was so computationally intensive. Ghez joined a Caltech group that was developing a speckle imager capable of working at infrared wavelengths emitted from interstellar-dust-shrouded objects such as active galaxies: spirals and ellipticals that had dusty but exceptionally bright regions at their centres. The thinking was that both the dust and the light emissions came from vast quantities of stars and gas spiralling into black holes millions or billions of times the mass of the Sun. Ghez was given the job of helping to write image-analysis software to give the speckle-imaging device the highest possible resolution. “Oh, this sounds great,” she says, remembering her reaction. “Black holes. Technique. I’m good.”

In the end, however, the imager couldn’t quite see the nuclei of active galaxies. So for her PhD dissertation, Ghez turned to brighter targets: newborn stars in our Galaxy. Astronomers knew in broad terms that stars are born in thick, interstellar gas clouds, when gravity pulls the gas into hot, dense knots that ignite with thermonuclear fusion. But they

**“That’s the evidence for a black hole. There are no alternatives we know of.”**

didn’t understand why a substantial fraction of the stars in the Sun’s neighbourhood are binaries: pairs of stars that orbit one another, often at close range. Were binaries born that way? Or did they somehow pair up later in life? Theory was not much help; the available star-formation models focused on single stars such as the Sun.

Ghez wanted to resolve the issue with data: do binary stars form as binaries, yes or no? (She had developed a taste for questions that could be answered crisply. “If you ask mushy questions,” she says, “you usually get mushy answers.”) Beginning in 1990, Ghez used her speckle-imaging techniques in two known star-forming regions to survey T Tauri objects — Sun-like stars just beginning their lives. With the higher resolution, she could see that many of the infant stars were indeed binaries, too young and too close together for the companions to have formed separately and then coalesced<sup>3</sup>. Ghez’s finding continues to have implications for issues such as the search for extrasolar planets, because the complex gravitational fields in a binary system are thought to make it difficult for planets to form in the vicinity.

Ghez focused on this topic exclusively until 1994, when she was hired at the University of California, Los Angeles, and gained access to a 10-metre instrument opened only a few years before at the Keck Observatory on Mauna Kea in Hawaii, and jointly owned by Caltech and the University of California. The biggest of a new generation of very big telescopes, Keck was ideal for catching the light of faint objects. But like all telescopes on the ground, its resolution was limited to what the atmosphere allowed. So Ghez applied her expertise in speckle imaging, realizing that its high resolution, combined with the telescope’s prodigious light-gathering power, would finally let her look at black holes.

Astronomers had long suspected that a supermassive black hole lay at the centre of every galaxy; active galaxies were the ones that just happened to have above-average quantities of matter feeding the black hole. If that



was true, a giant black hole ought to lie at the centre of the Milky Way. That region is difficult to observe from Earth because there is a lot of interstellar gas and dust in the way and only certain wavelengths of radiation can make it through. But observations in  $\gamma$ -rays, X-rays and radio and infrared emissions were consistent with the presence of gas moving at high velocities, and stars crowding together in high numbers. They had to be orbiting something massive and hard to see.

But what? To remove any doubt about the nature of the central object, observers would have to show that it was too small to be anything but a black hole. (The fundamental equations of gravity guarantee that, for a given mass, a black hole is smaller than any possible system made of ordinary matter.)

Ghez and her team decided to tackle the problem with a straightforward but tedious strategy: track stars for years and decades as they orbit the central object. The radius and period of each orbit would give the central object's mass, and the distance of closest approach would put an upper limit on its size. The astronomers started in 1995 by mapping the positions of stars in a dense cluster near the suspected location of the black hole. Every year or so, they mapped the positions again.

"With two points you draw a line and get the velocity," says Ghez. "With three points, you actually believe your line." And then you keep going. "You wait long enough, keep taking pictures" and you should be able to start tracing out the curves of the stars' orbits. By 1999, Ghez and her colleagues had done exactly that with three stars<sup>4</sup>.

## SEEING THE LIGHT

These observations showed the orbit in only two dimensions, projected onto the plane of the sky. Getting the full orbit required the third dimension, the star's motion towards and away from Earth, which meant measuring the spectrum of the light it emitted and watching how its velocity shifted the wavelength of the spectral lines — measurements that required more light than speckle imaging could supply. Ghez, however, was already working with astronomers developing a more advanced technology called adaptive optics. The idea was to take the light gathered by the telescope and bounce it off a flexible mirror that could be continuously deformed to counteract distortions caused by the atmosphere. The technique wasted much less light than speckle imaging did, so it promised higher resolution and the ability to capture the spectra of fainter objects. Ghez was doing proof-of-concept observations, and was "right there and ready to snatch it up the minute it was ready to go", she says.

In June 2002, Ghez and her team used adaptive optics to observe one of their three stars, called S0-2, and found that by adding in older speckle data they could map its complete orbit<sup>1</sup>. The orbit took 16 years, implied a central mass of 4.1 million Suns and showed that the star's closest approach to that mass was less than twice the distance from Pluto to the Sun — meaning that the central object was no larger than the Solar System. With that much mass in that small a space, says Ghez, "that's the evidence for a black hole. There are no alternatives that we know of."

That result was satisfying but not unexpected. Much more surprising was that S0-2's spectrum showed it to be less than 10 million years old. No star orbiting close to a black hole should be that young, because the black hole's tidal forces would shred any cloud of gas and dust before it could form. There was no good explanation for why the young stars



Andrea Ghez wants to find out what lies at the heart of our Galaxy.

are there, says Thomas Prince, an astronomer at Caltech and Ghez's first adviser, "and there still is not".

In subsequent years, as adaptive optics improved, Ghez and her team measured the velocities of thousands of stars in the Galactic Centre, and estimated the orbits of about two dozen; the latest is a star named S0-102, which has the shortest orbit yet at 11.5 years (ref. 2). Most of them also proved to be equally young and enigmatic. "It's beautiful work really," says Rosemary Wyse, an astronomer at Johns Hopkins University in Baltimore, Maryland, who studies the Milky Way. "It's one of those painstakingly technical tours de force that produces good science results."

Only one other team in the world is doing this type of work: Genzel's. Both groups have access to large telescopes and sophisticated adaptive-optics systems, and the word is that competition is fierce — which neither Ghez nor Genzel denies, although they try to play it down. Perhaps more important is that they agree on the same measurements on the same stars, differing mainly in interpretations.

The latest example concerns an object that might help to show how a black hole destroys its prey. Both Ghez and Genzel agree that it is hot, red and heading straight towards the Galactic Centre. Genzel says that it is a cloud of gas with the mass of three Earths, that it will be torn to smithereens by the black hole and that in the process it will radiate X-rays fiercely<sup>3</sup>. Ghez maintains that it is hard to say what the object is; it could equally be a star moving through gas, in which case the black hole is unlikely to disrupt it. Genzel says that the object's closest approach to the black hole will be this autumn; Ghez says maybe later and in any case the timing is uncertain. Avi Loeb, an astronomer at Harvard University in Cambridge, Massachusetts, agrees with both Ghez and Genzel that this kind of competition over difficult, expensive observations makes the resulting astronomy more credible. "Overall," he says, "it's good for science."

Ghez is already looking to the next technologies she can adopt. She is on the scientific advisory committee of the Thirty Meter Telescope, which is due to start observations from Mauna Kea towards the end of the decade and will gather roughly ten times more light than Keck. She is also advising the developers of a version of adaptive optics that promises even higher resolutions than are currently available.

"I'm positioning myself," says Ghez. "I want to know how to use these systems, I want to know how to use the data, I want this all to work." Like other great scientists, says Thomas Soifer, an astronomer at Caltech and another of her early advisers, "she has this single-minded attitude of, 'I'm going to beat this problem into submission'".

Ghez is 47 years old and the stars that she has been tracking since 1995 have orbits ranging from tens to hundreds of years. How long does she plan on living? "I don't know," says Ghez. "I figure the more fun I have, the longer I'll live." ■

*Ann Finkbeiner is a freelance writer in Baltimore, Maryland.*

1. Ghez, A. M. *et al. Astrophys. J.* **586**, L127–L131 (2003).
2. Meyer, L. *et al. Science* **338**, 84–87 (2012).
3. Ghez, A. M., Neugebauer, G. & Matthews, K. *Astron. J.* **106**, 2005–2023 (1993).
4. Ghez, A. M., Morris, M., Becklin, E. E., Tanner, A. & Kremenek, T. *Nature* **407**, 349–351 (2000).
5. Gillessen, S. *et al. Nature* **481**, 51–54 (2012).

# SLOW SCIENCE

*The world's longest-running experiments remind us that science is a marathon, not a sprint.*

BY BRIAN OWENS

**A**lthough science is a long-term pursuit, research is often practised over short timescales: a discrete experiment or a self-contained project constrained by the length of a funding cycle. But some investigations cannot be rushed. To study human lifespans or the roiling of Earth's crust and the Sun's surface, for instance, requires decades and even centuries.

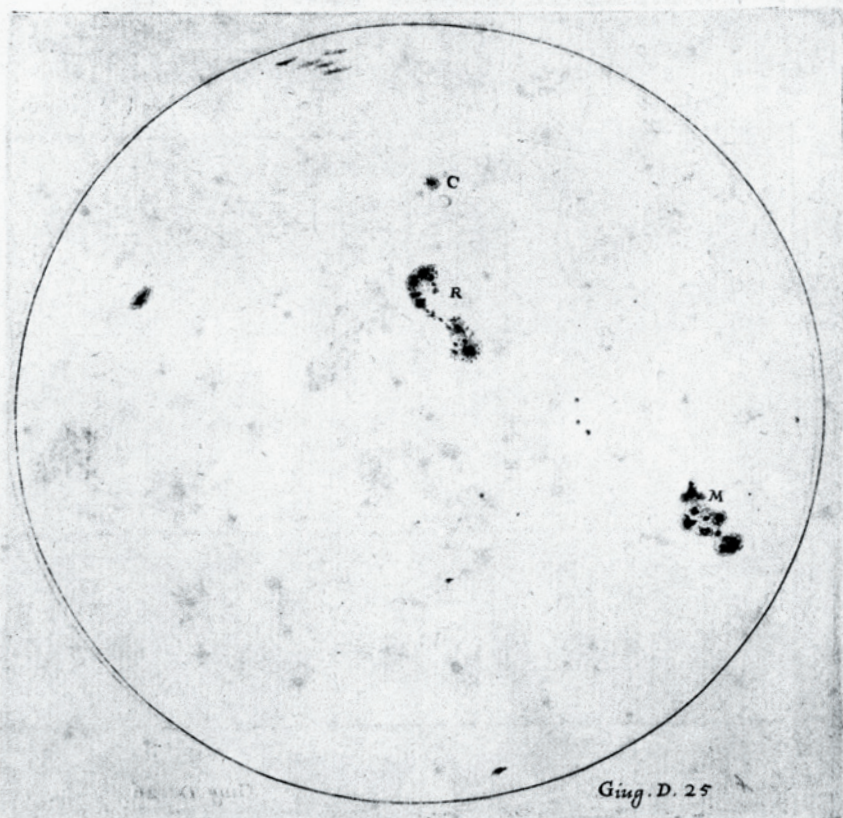
Here, *Nature* takes a look at five of science's longest-running projects, some of which have been amassing data continuously for centuries. Some generate hundreds of papers a year; one produces a single data point per decade.

Experiments operating at this pace are challenged by shifting research priorities and technologies, and their existence is regularly threatened by funding droughts and changes in stewardship. But they are bound together by the foresight of the scientists who started them and the

patience and dedication of those who carry the torch. If persistence predicts a long and healthy life — as one 90-year study of human longevity has suggested — then the scientists featured here could set some records themselves.

➔ **NATURE.COM**

To hear a podcast about these experiments, visit: [go.nature.com/jlada7](http://go.nature.com/jlada7)



Galileo was drawing sunspots as early as 1613.

## 400 YEARS **COUNTING SPOTS**

Astronomers have been recording the appearance of sunspots ever since the telescope was invented more than 400 years ago; even Galileo recorded his observations. But early observers had no knowledge of what the dark patches on the Sun's surface were, or of the magnetic fields that created them. That began to change when, in 1848, the Swiss astronomer Rudolf Wolf began making systematic observations and developed a formula that is still used today to calculate the international sunspot number, also known as the Wolf number, which gives a measure of how solar activity is changing over time.

In 2011, Frédéric Clette became director of the Solar Influences Data Analysis Center, based at the Royal Observatory of Belgium in Uccle, which curates sunspot counts gleaned from photographs and hand drawings of the Sun's surface made by more than 500 observers since 1700.

The data are invaluable for predicting sunspot activity, says Leif Svalgaard, a solar physicist at Stanford University in California. The activity seems to wax and wane over the course of 11 years or so, and the streams of charged particles that the sunspots spray into space can affect satellites and electronics on Earth. The detailed records help researchers to understand why that cycle happens, and to refine predictions of particularly intense events. "The longer the time series is, the better we can check our theories," Svalgaard says. Around 200 papers a year cite sunspot data, in fields extending beyond solar physics to geomagnetism, atmospheric science and climate science.

But the enterprise runs largely on goodwill. Each month, the Belgian centre collates sunspot numbers from about 90 observers, two-thirds of them amateurs, who use small optical telescopes no more powerful than those available 200 years ago. And although it is a World Data

UNIVERSAL HISTORY ARCHIVE/UGC/BRIDGEMAN ART LIBR.



Centre recognized by the International Council for Science in Paris, it receives no funding from the organization. Clette works with one other part-time person to maintain the database, in addition to his 'night-job' as an astronomer at the Royal Observatory of Belgium.

Still, says Clette, it is fascinating to 'work' with colleagues from hundreds of years ago. For instance, he says that even though Galileo's coverage of the Sun was spotty because Galileo was "busy with planets and other things", the drawings are detailed enough to reveal information about the magnetic structure of the sunspot groups and the size and tilt of the star's dipole. "You can extract from those drawings exactly the same information as from a drawing made today," he says.

More than that, however, he is taken with his forebears' foresight. They faithfully recorded what they saw, thinking that it could be useful later on, he says. "It's a fundamental aspect of science," he says, "not worrying what will be the final result."

## 170 YEARS **MONITORING AN IRRITABLE GIANT**

Although consistently active, every few thousand years, Mount Vesuvius erupts in spectacular style. The last time it did so, in AD 79, it consumed the city of Pompeii in the flames and before that, about 3,800 years ago, it covered all of present-day Naples in hot gas and rock (see *Nature* 473, 140–141; 2011). The Vesuvius Observatory, the oldest volcano research station in the world, has been keeping an eye on its inhospitable subject since 1841, logging all the volcano's seismic rumbles to try to spot approaching danger. Originally perched 600 metres up the side of the volcano, far enough from the summit to be safe from ejected debris and high enough on a knoll to avoid the lava flows, the observatory has shaped the way that volcanology and

geology is done, says Marcello Martini, its current director.

Macedonio Melloni, the observatory's first director, did pioneering work on the magnetic properties of lava that was crucial to later studies of palaeomagnetism — the history of Earth's magnetic field as recorded in rocks. In 1856, its second director, Luigi Palmieri, invented the electromagnetic seismograph, which was much more sensitive to ground tremors than previous machines and allowed him to predict eruptions. Under Palmieri and subsequent directors the observatory contributed to the development of much of the instrumentation used to monitor volcanoes worldwide. In the early twentieth century, for example, Giuseppe Mercalli developed the scale still used today to classify volcanic activity.

But the building itself no longer has the same role. "In the early stages it was important to be as close to the action as possible, but that's not necessarily the case any more," says Haraldur Sigurdsson, a volcanologist at the University of Rhode Island in Kingston. Most of the monitoring is now done by remote, ground-based sensors, with the data sent back to the lab of the National Institute of Geophysics and Volcanology in Naples. The original buildings were turned into a museum in 1970.

In addition to informing scientific theory, the observations are used to predict trouble and protect the public — as they did successfully in 1944. The Naples lab, where scientists are on duty 24 hours a day, also keeps an eye on Mount Stromboli, on an island north of Sicily; the Campi Flegrei caldera west of Naples; and the island of Ischia. Sigurdsson, says, however, that the future of volcanology lies not with putting sensors on the volcanoes already known to be dangerous, but with satellite-based radar that can study ground deformation everywhere and pick out the risky regions irrespective of geologists' expectations. "We should be moving towards an internationally coordinated system of volcano monitoring that is not tied down to bricks and mortar on the side of a volcano, but looking at it in a really comprehensive way globally," he says.

FRATELLI ALINARI MUS. COLLECTIONS, FLORENCE



Now a museum, the original building of the Vesuvius Observatory was used to monitor volcanic activity from the side of the mountain.





ROTHAMSTED RESEARCH

Rothamsted has been home to experiments on the effects of fertilizer on the yield of wheat since 1843.

## 170 YEARS HARVESTING DATA

The stewards of long-term research projects are keen to maintain the integrity of the work, but also to keep it relevant. That is the case for Andy Macdonald who, in 2008, inherited a set of agricultural experiments that have been testing the effects of mineral fertilizers and organic manure on crop production since 1843.

Started by the fertilizer magnate John Lawes on the grounds of his country estate at Rothamsted, north of London, these studies have been used to test how nitrogen, phosphorus, potassium, sodium, magnesium and farmyard manure affect the yields of several staple crops, including wheat, barley, legumes and root crops.

"After 20 or 30 years, many of the basic questions about the relative importance of different fertilizers were pretty well answered," says Macdonald, manager of the 'classical experiments' now run at Rothamsted Research. Nitrogen has the largest effect, followed by phosphorus. So the experiments are periodically updated to test new ideas and keep them relevant to current farming practice. In 1968, for example, the long-strawed cereal crops that had been grown since the experiments started were replaced by the higher-yielding short-strawed cereal crops being adopted by farmers. Macdonald says that these new crops turned out to need more fertilizer than the traditional cultivars because of the additional nutrients they were extracting from the soil, so farmers had to adapt.

"Rothamsted is the granddaddy of long-term agricultural research," says Phil Robertson, director of the W. K. Kellogg Biological Station, a long-term agricultural research site at Michigan State University in Hickory Corners. The unbroken chain of data is invaluable, he says. Not only is Rothamsted able to study environmental and biological trends — such as carbon storage in soil or the effects of invasive species — that become apparent only over long timescales, but it also provides a platform for shorter studies on, say, nitrate loss in soil.

The Rothamsted archive holds about 300,000 preserved plants and soil samples that have been collected since the experiments began. In 2003, scientists extracted the DNA of two wheat pathogens from archived samples dating back to 1843 and showed that industrial sulphur dioxide emissions affected which one was dominant<sup>1</sup>.

It can be a struggle to keep funding agencies interested. Rothamsted gets by on a mixture of government funding, competitive grants and a trust fund set up by Lawes before he died. "As a funder you have to

commit to maintaining observations even during periods where there may not be any exciting results," says Robertson, who was involved in setting up the US Department of Agriculture's Long Term Agro-Ecosystem Research network last year. Macdonald and his team take pride in their history. "I think back sometimes to John Lawes," Macdonald says. "I feel a great responsibility to ensure that the experiments are handed on in good condition for the next generations. They're not a museum piece, they're part of our living scientific community."



STANFORD UNIV. ARCHIVE

Lewis Terman started one of the lengthiest studies of human development.

## 90 YEARS WATCHING GENIUS BLOSSOM

In 1921, psychologist Lewis Terman of Stanford University in California started tracking more than 1,500 gifted children — as identified by the Stanford-Binet IQ test that he developed — born between 1900 and 1925. It was one of the world's first longitudinal studies and it has now accrued the longest in-depth record of human development, having tracked the participants for nine decades — looking at their home lives, education, interests, abilities and personality.

One of Terman's main goals in his 'Genetic Studies of Genius' was to



disprove the then-common assumption that gifted children were sickly, socially inept and not well-rounded. But even by the standards of his day, the study design was plagued with problems. His selection method was haphazard, administering the tests based largely on recommendations from teachers, and the sample was far from representative (more than 90% were white and upper or middle class, and Terman even enrolled his own children). What is more, Terman skewed the very life outcomes he was trying to study, writing letters of recommendation for his 'Termites', as the participants became known, and helping several to get into Stanford.

By tracking the children into adulthood, Terman showed that they were just as healthy and well-adjusted as the general population and that they generally grew into successful, happy adults. And, as it progressed, researchers tweaked the study to try to overcome some of its flaws.

In the 1980s, for example, George Vaillant, a psychologist at Harvard Medical School in Boston, Massachusetts, began using Terman data to supplement his own long-running study of adult development and started collecting the death certificates of the Termites. Using these records, Howard Friedman, a psychologist at the University of California, Riverside, was able to come up with one of the Terman study's most significant findings. He showed that conscientiousness — namely prudence, persistence and planning — measured in both childhood and adulthood is a key psychological factor predicting longevity, and was associated with as many as six or seven extra years of life<sup>2</sup>. "It would not have been easy to discover this without a lifespan longitudinal prospective data set," says Friedman.

Longitudinal studies evolve according to scientific fashions, says Laura Carstensen, director of the Stanford Center on Longevity. New investigators will add measures and modify or discard those that they think are no longer interesting or are obsolete. "Today we would measure emotional well-being, for example, in a very different way from in 1900," she says. So in many ways "looking at a longitudinal data set is almost like writing a history of psychology".

## 85 YEARS WAITING FOR THE DROP

On his second day of work at the University of Queensland in Brisbane, Australia, in 1961, physicist John Mainstone came across a quirky little experiment that had been quietly running in a cupboard for 34 years. Fifty years later, he is still looking after it — and still waiting to witness its most dramatic activity.

The pitch-drop experiment started when Thomas Parnell, the university's first professor of physics, set up a demonstration for his students to show that a sample of pitch, a black tar distillate that is brittle enough to shatter with a hammer when cold, will act like a liquid and flow through a funnel, dripping out of the bottom like the world's slowest hour glass. It did, at a rate of about one drop every 6–12 years. Mainstone expects — cautiously — the ninth drop to fall sometime towards the end of this year.

The experiment is not exactly a hotbed of discovery. In 86 years, it has yielded exactly one scientific paper<sup>3</sup>, which calculated that the pitch was 230 billion times more viscous than water. And in 2005, it earned the dubious distinction of an Ig Nobel prize (see *Nature* 437, 938–939; 2005), a cheeky parody of the Nobels.

There is, however, still some science to be gleaned. No one has ever seen a drop fall — a webcam recording the experiment failed just when the most recent drop fell, in November 2000 — so exactly what happens when the drop detaches from the body of pitch above is unknown. It will also take another few decades to tease out how weather, the introduction of air conditioning and the vibrations from renovation work in the building influence the drip rate.

But Mainstone says that the experiment's value lies not in its science,

"IT'S GOING ABOUT ITS BUSINESS WHILE THE WORLD IS GOING THROUGH ALL SORTS OF TURMOIL."



The pitch-drop experiment has yielded just eight drops since 1927.

but in its historical and cultural impact: it has inspired sculptors, poets and writers to muse about the passage of time and the pace of modern life. It also provides a link to scientific history, and a sense of constancy. "It's going about its business while the world is going through all sorts of turmoil," Mainstone says. And with a great deal of pitch left in the funnel, it has the potential to serenely ignore that turmoil for another 150 years or so. Luckily, 78-year-old Mainstone has already convinced a younger colleague to watch over the experiment after he is gone. ■

**Brian Owens** is a freelance writer in New Brunswick, Canada.

1. Bearchell, S. J., Fraaije, B. A., Shaw, M. W. & Fitt, B. D. *Proc. Natl Acad. Sci. USA* **102**, 5438–5442 (2005).
2. Friedman, H. S. & Martin, L. R. *The Longevity Project: Surprising Discoveries for Health and Long Life from the Landmark Eight-Decade Study* (Hudson Street Press, 2011).
3. Edgeworth, R., Dalton, B. J. & Parnell, T. *Eur. J. Phys.* **5**, 198–200 (1984).

# COMMENT

**ENVIRONMENT** Conservationists call for a global zoning exercise for roads **p.308**



**HISTORY** Ripping yarn of the ape-man of Victorian England **p.310**

**EVOLUTION** First biography of W.D. Hamilton, the gentle giant of genetics **p.313**

**FUNDING** Australia's grant system wastes centuries of researchers' time **p.314**

ILLUSTRATION BY PHIL DISLEY



## Sustainable development goals for people and planet

Planetary stability must be integrated with United Nations targets to fight poverty and secure human well-being, argue **David Griggs** and colleagues.

**T**he United Nations Rio+20 summit in Brazil in 2012 committed governments to create a set of sustainable development goals (SDGs) that would be integrated into the follow-up to the Millennium Development Goals (MDGs) after their 2015 deadline. Discussions on how to formulate these continue this week at UN headquarters in New York.

We argue that the protection of Earth's

life-support system and poverty reduction must be the twin priorities for SDGs. It is not enough simply to extend MDGs, as some are suggesting, because humans are transforming the planet in ways that could undermine development gains.

As mounting research shows, the stable functioning of Earth systems — including the atmosphere, oceans, forests, waterways, biodiversity and biogeochemical cycles — is

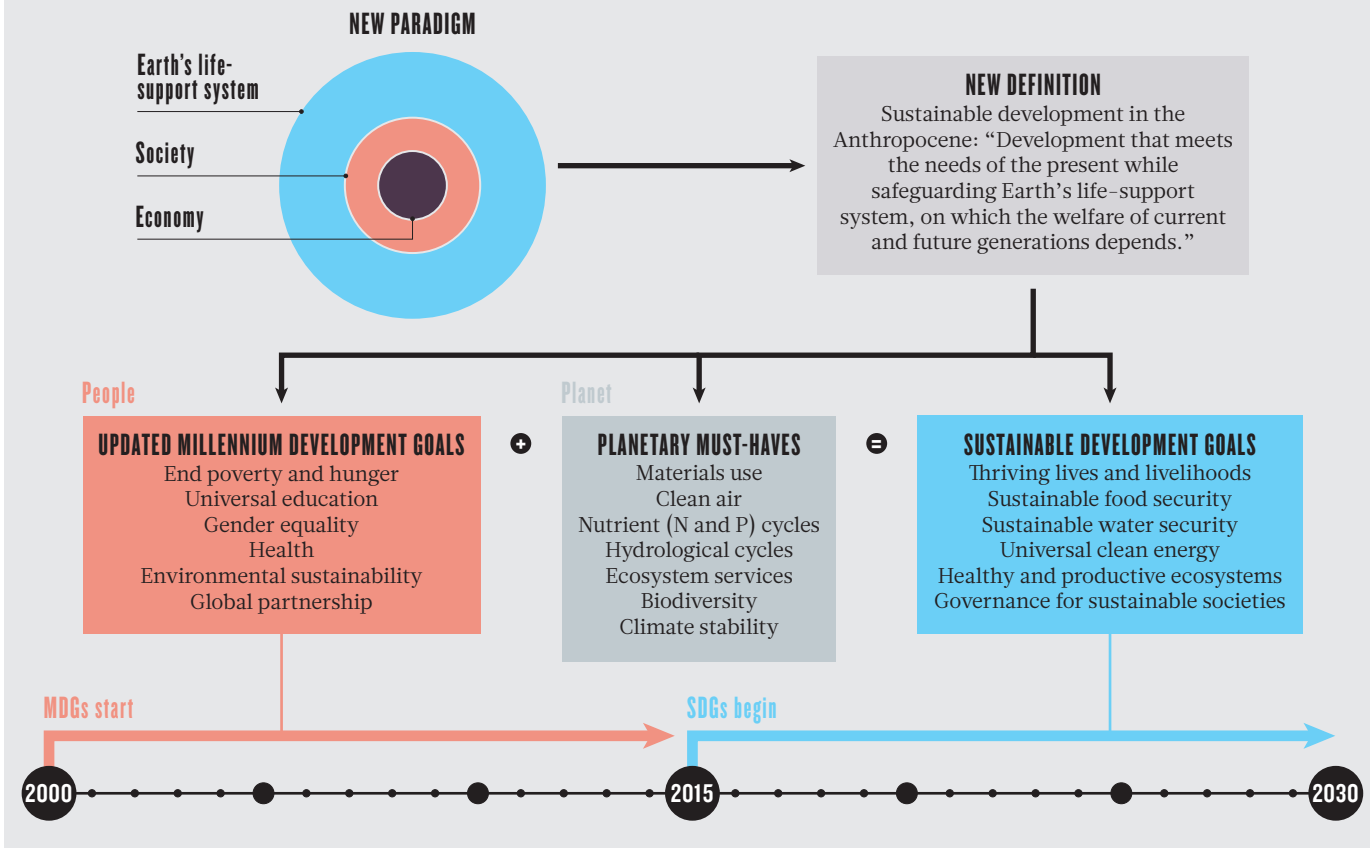
a prerequisite for a thriving global society. With the human population set to rise to 9 billion by 2050, definitions of sustainable development must be revised to include the security of people and the planet.

Defining a unified set of SDGs is challenging, especially when there can be conflict between individual goals, such as energy provision and climate-change prevention. But we show here that it is possible. By ►



# A UNIFIED FRAMEWORK

A set of six sustainable development goals (SDGs) follow from combining the Millennium Development Goals (MDGs) with conditions necessary to assure the stability of Earth's systems.



► combining the MDGs with global environmental targets drawn from science and from existing international agreements, we propose six SDGs with provisional targets for 2030.

## ENTER THE ANTHROPOCENE

Since 2000, the MDGs have focused on reducing extreme poverty in developing countries. But pursuing a post-2015 agenda focused only on poverty alleviation could undermine the agenda's purpose. Growing evidence and real-world changes convincingly show that humanity is driving global environmental change and has pushed us into a new geological epoch — the Anthropocene<sup>1</sup>.

Further human pressure risks causing widespread, abrupt and possibly irreversible changes to basic Earth-system processes. Water shortages, extreme weather, deteriorating conditions for food production, ecosystem loss, ocean acidification and sea-level rise are real dangers that could threaten development and trigger humanitarian crises across the globe. Growing affluence and the right to development among the world's poor demand that

people of all nations make the transition to sustainable lifestyles.

By coordinating actions internationally, SDGs can address these risks. The MDGs have shown that a goal-setting approach raises both public and policy support and channels funds effectively towards urgent global problems<sup>2</sup>. However, the political reluctance to go beyond merely extending the MDGs is a concern.

The targets for the SDGs must be measurable, based on the latest research and should apply to developed and developing countries. First, however, we need to reframe the UN paradigm of three pillars of sustainable development — economic, social and environmental — and instead view it as a nested concept. The global economy services society, which lies within Earth's life-support system. The definition of sustainable development, as laid out in the 1987 report from the UN World Commission on Environment and Development (the Brundtland Commission), should therefore be redefined to "development that meets the needs of the present while safeguarding Earth's life-support system, on which the welfare

of current and future generations depends".

To set appropriate goals and targets, environmental conditions have to be identified that enable prosperous human development and set tolerable ranges for the biosphere to remain in that state. The extraordinarily stable Holocene epoch that allowed our ancestors to develop agriculture and modern societies during the past 10,000 years provides a scientific reference point. Indeed, these are the only conditions we know that can support modern life.

Building on decades of research, a 2009 analysis defined planetary boundaries which would be unsafe to transgress for nine Earth-system processes<sup>3</sup>: climate change; rate of biodiversity loss (terrestrial and marine); interference with the nitrogen and phosphorus cycles; stratospheric ozone depletion; ocean acidification; global freshwater use; change in land use; chemical pollution; and atmospheric aerosol loading.

Adapting this planetary boundaries work, and using recent credible scientific studies and existing international processes — such as the United Nations Framework Convention on Climate Change — we extracted a

## SIX SUSTAINABLE DEVELOPMENT GOALS

## Some provisional targets for 2030

● **Goal 1: Thriving lives and livelihoods. End poverty and improve well-being through access to education, employment and information, better health and housing, and reduced inequality while moving towards sustainable consumption and production.**

This extends many targets of the Millennium Development Goals (MDGs) on poverty, health and urban environments and applies them to developed nations. It should include targets on clean air that build on World Health Organization guidelines for pollutants such as black carbon<sup>6</sup>; reductions in emissions of stratospheric ozone-depleting substances in line with projections from the Montreal Protocol; critical loads for man-made chemical compounds and toxic materials; and sustainable practices for extraction, use and recycling of scarce minerals and metals and other natural resources.

● **Goal 2: Sustainable food security. End hunger and achieve long-term food security — including better nutrition — through sustainable systems of production, distribution and consumption.**

The MDG hunger target should be extended and targets added to limit nitrogen and phosphorus use in agriculture<sup>3,7,8</sup>. Nutrient-use efficiency should improve by 20% by 2020; no more than 35 million tonnes of nitrogen per year should be extracted from the

atmosphere; phosphorus flow to the oceans should not exceed 10 million tonnes a year; and phosphorus runoff to lakes and rivers should halve by 2030.

● **Goal 3: Sustainable water security. Achieve universal access to clean water and basic sanitation, and ensure efficient allocation through integrated water-resource management.**

This would contribute to MDG health targets, restrict global water runoff to less than 4000 cubic kilometres a year and limit volumes withdrawn from river basins to no more than 50–80% of mean annual flow<sup>3,9</sup>.

● **Goal 4: Universal clean energy. Improve universal, affordable access to clean energy that minimizes local pollution and health impacts and mitigates global warming.**

This contributes to the UN commitment to sustainable energy for all, and addresses MDG targets on education, gender equity and health. To ensure at least a 50% probability of staying within 2°C warming<sup>10</sup>, sustainability targets should aim for global greenhouse-gas emissions to peak in 2015–20, drop by 3–5% a year until 2030, and fall by 50–80% by 2050.

● **Goal 5: Healthy and productive ecosystems. Sustain biodiversity and**

**ecosystem services through better management, valuation, measurement, conservation and restoration.**

This combines the MDG environmental targets with 2030 projections of the Aichi Targets adopted by the Convention on Biological Diversity (see [www.cbd.int/sp/targets](http://www.cbd.int/sp/targets)). Extinctions should not exceed ten times the natural background rate. At least 70% of species in any ecosystem and 70% of forests should be retained. Aquatic and marine ecosystems should be managed to safeguard areas crucial for biodiversity, ecosystem services and fisheries.

● **Goal 6: Governance for sustainable societies. Transform governance and institutions at all levels to address the other five sustainable development goals.**

This would build on MDG partnerships and incorporate environmental and social targets into global trade, investment and finance<sup>4</sup>. Subsidies on fossil fuels and policies that support unsustainable agricultural and fisheries practices should be eliminated by 2020; product prices should incorporate social and environmental impacts. National monitoring, reporting and verification systems must be established for sustainable-development targets; and open access to information and decision-making processes should be secured at all levels.

list of sustainability ‘must-haves’ for human prosperity (see ‘A unified framework’).

We combined these with the MDG targets, updated and extended for 2030, to produce six SDGs: thriving lives and livelihoods, sustainable food security, sustainable water security, universal clean energy, healthy and productive ecosystems, and governance for sustainable societies (see ‘Some provisional targets for 2030’).

The driving principles remain: reducing poverty and hunger, improving health and well-being and creating sustainable production and consumption patterns. A goal of improving lives and livelihoods, for example, would promote sustainable access to food, water and energy while protecting biodiversity and ecosystem services.

None of this is possible without changes to the economic playing field<sup>4</sup>. National policies should, like carbon pricing, place a value on natural capital and a cost on unsustainable actions. International governance of the global commons should be strengthened, for example through binding agreements on climate change, by halting the loss of biodiversity and ecosystem services and

by addressing other sustainability concerns.

The SDG framework manages trade-offs and maximizes synergies between targets, and can be implemented from international to city scales. It integrates social, economic and environmental dimensions and provides guidance for humanity to prosper in the long term. A small number of

**“None of this is possible without changes to the economic playing field.”**

goals is essential for focus; others could be added but should build on the core six.

There are many gaps and uncertainties in our knowledge of global environmental risks and how to enable societies to become resource-efficient, sustainable and wealthy. Research initiatives such as Future Earth, a ten-year programme coordinated by the International Council for Science<sup>5</sup>, are needed to refine targets and provide sustainable solutions for human well-being.

But the first step is for policy-makers to embrace a unified environmental and social framework for the SDGs, so that today’s advances in development are not lost as our

planet ceases to function for the benefit of a global population. ■

**David Griggs** is professor and director of the Monash Sustainability Institute, Monash University, Victoria 3800, Australia. **Mark Stafford-Smith, Owen Gaffney, Johan Rockström, Marcus C. Öhman, Priya Shyamsundar, Will Steffen, Gisbert Glaser, Norichika Kanie, Ian Noble.** e-mail: [dave.griggs@monash.edu](mailto:dave.griggs@monash.edu)

1. Steffen, W. *et al.* *Ambio* **40**, 739–761 (2011).
2. United Nations. *The Millennium Development Goals Report 2012* (UN, 2012).
3. Rockström, J. *et al.* *Nature* **461**, 472–475 (2009).
4. Biermann, F. *et al.* *Science* **335**, 1306–1307 (2012).
5. Glaser, G. *Nature* **491**, 35 (2012).
6. Shindell, D. *et al.* *Science* **335**, 183–189 (2012).
7. Sutton, M. A. *et al.* *Our Nutrient World* (Centre for Ecology and Hydrology, 2013).
8. Carpenter, S. R. & Bennett, E. M. *Environ. Res. Lett.* **6**, 014009 (2011).
9. Smakhtin, V. U. & Batchelor, A. L. *Hydrol. Process.* **19**, 1293–1305 (2005).
10. Huntingford, C. *et al.* *Environ. Res. Lett.* **7**, 014039 (2012).

Supplementary information and full author affiliations accompany this article online at [go.nature.com/zowqiw](http://go.nature.com/zowqiw).





ANDRÉ VEIRA/POLARIS/EYEVINE

Trucks from an illegal logging operation transport timber along a partially paved highway that cuts deep into the Brazilian Amazon.

# A global map for road building

Roads are proliferating across the planet. Located and designed wisely, they can help rather than harm the environment, argue **William F. Laurance** and **Andrew Balmford**.

Nearly 100,000 kilometres of road criss-cross the Amazon rainforest. That is enough to circle the Earth two-and-a-half times. Even in formerly remote corners of the world — the Congo region, Borneo, Siberia, Namibia — road networks are expanding apace. This global road rush is being driven by escalating demand for minerals, fossil fuels, timber and arable land, and by developing nations as they work to improve transport and energy infrastructures<sup>1</sup>.

Roads can bring myriad environmental problems. In the Amazon, new roads in forested areas often promote illegal colonization of undisturbed areas, as well as mining, hunting and land speculation. More than 95% of deforestation, fires and atmospheric carbon emissions in the Brazilian Amazon occur within 50 kilometres of a road<sup>2</sup>.

Yet the effects of a road vary depending on its location and design. A paved highway slicing through a large forest tract

can precipitate an environmental disaster. Conversely, in places where farming is already widespread and intact habitat is scarce, and where there are sizeable gaps between current and potential farm yields, building high-quality roads can improve farms' efficiency, increase their profitability and limit their environmental impact.

We propose that environmental scientists, planners, road engineers and other stakeholders carry out a global 'road-zoning' project to map areas that should remain road-free and those in which transport urgently needs improving.

## LAND-USE PRESSURES

The twenty-first century will bring profound changes in land use, many of them unavoidable and even desirable. Food demand is projected to double by 2050. Under current farming practices, this would require an additional 1 billion hectares of farming and grazing land<sup>3</sup> — an area the size of Canada.

Given escalating demand for food, fibre and biofuels, researchers and policy-makers have focused on improving agriculture through the use of modern crop varieties, fertilizers, pest control and better transport. The hope is that such technologies will allow farmers to increase yields without using too much extra land<sup>4</sup>.

In practice, however, by making farming more profitable, yield improvements can encourage the conversion of land for more crop and livestock production<sup>5</sup>. For instance, as a result of high demand and innovations such as new cultivars, oil-palm plantations are rapidly expanding across the tropics — often at the expense of biodiversity-rich rainforests.

We are convinced that increasing agricultural yields will lessen the impact of farming on natural ecosystems only if coupled with effective land-use planning<sup>6</sup>. Roads, which profoundly influence the footprint of human activities, are a key element of such planning.

Numerous factors — economic, political, social and practical — influence road

planning. But a few key principles could help to guide the location and design of roads. For example, the most environmentally damaging roads are those that penetrate relatively pristine regions. This is because the probability that a patch of land will be cleared rises dramatically if an adjacent area has already been cleared<sup>7</sup>. For this reason, the first cut into a forest is also the most crucial.

Furthermore, paved highways typically have much larger environmental impacts than unpaved roads, especially in wet environments in which unpaved roads can become seasonally impassable. In Brazil, for instance, the Belém-Brasília Highway, completed in the early 1970s, now cuts a 400-kilometre-wide swathe of cleared forest and secondary roads through the Amazon<sup>2</sup>.

In certain contexts, however, road building, or improvements such as paving, can be socially and environmentally beneficial. Often, agriculture follows roads created for other purposes, such as mining or logging. This can result in the expansion of farms into places with marginal soils or climates, or into locations that are too far from markets to be cost-effective. Conversely, well-planned roads can increase farmers' access to markets, reducing waste and improving profits. Anecdotal evidence indicates that ongoing road improvements in parts of sub-Saharan Africa are gradually raising rural farmers' access to fertilizers and increasing their capacity to transport crops to markets.

Several studies suggest that road improvements in areas suited to agricultural development can attract migrants away from vulnerable areas, such as the edges of pristine forests<sup>8</sup>. Concentrating people in carefully defined areas is beneficial, because the relationship between deforestation and population density is nonlinear. In general, migrants entering an already cleared area remove much less forest than those

who arrived first<sup>2</sup>; latecomers may include merchants instead of farmers and loggers, for instance, or farmers who buy up previously established plots.

It is much easier for policy-makers to influence patterns of road development than to affect more socially complex problems such as population growth and overconsumption. Roads can be re-routed, cancelled or delayed. Large road projects are often funded by taxpayers, investors or international donors who can be surprisingly responsive to environmental concerns. For example, if corporations that build environmentally bad roads are publicly named, they can lose customers and shareholders. Concord Pacific, a Malaysian logging corporation, was publicly castigated in the early 2000s for bulldozing a 180-kilometre-long road into the highlands of Papua New Guinea — ostensibly to aid local communities. After the company grabbed more than US\$60 million in illegal timber<sup>9</sup>, it was fined \$97 million by the national court of Papua New Guinea.

#### TRANSPORT PLANNING

We believe that a collaborative, global zoning exercise is needed to identify where road building or improvement should be a priority, where it should be restricted and where existing roads should be closed. A multidisciplinary team could integrate and standardize satellite data on intact habitats with information on transport infrastructure, agricultural yields and losses, biodiversity indicators, carbon storage and other relevant factors. Much of this information has been recorded or can be extrapolated from current data sources.

The next step would be to make the results of these analyses readily available as high-resolution, colour-coded maps to aid policy-makers, conservation organizations and others in planning roads. This mapping

exercise could be repeated, and at finer scales, as circumstances on the ground change or data sources improve over time.

Beyond limiting habitat destruction, a global road-zoning scheme could safeguard rare environments and areas rich in endemic species. It could prove useful for a number of projects already planned or under way, such as the Tanzanian government's plan (which is being legally challenged) to bisect Serengeti National Park with a highway that could disrupt the area's world-famous migration of wildebeest and other wildlife, or the Ladia Galaska road network under construction in northern Sumatra, which will cut through hundreds of kilometres of rainforest and protected areas when complete.

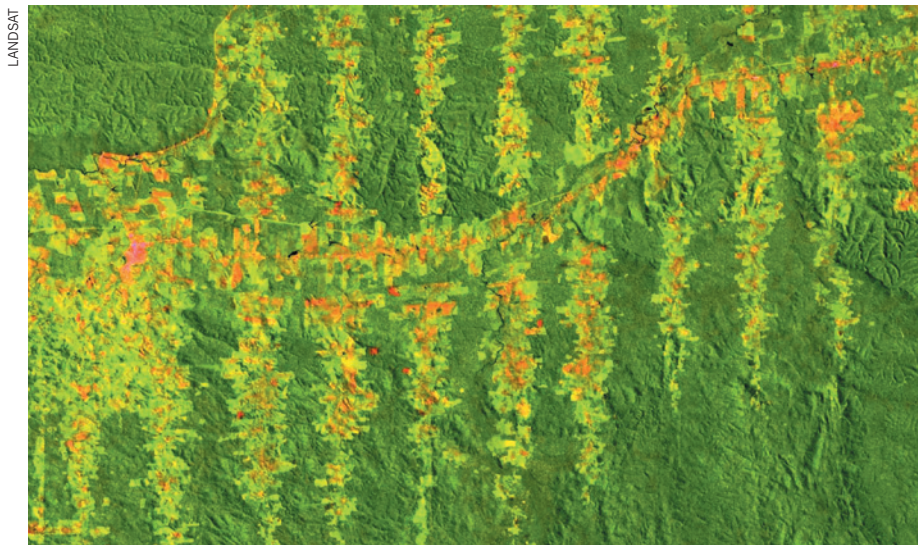
For transport projects that have high environmental costs but seem unavoidable, such as Brazil's Manaus-Porto Velho highway — which is now under construction and has the potential to speed settlers and land specu-

lators into the heart of the Amazon when complete — alternatives such as railways or river transport might be effective compromises. Trains and boats move people and products but limit the human footprint by stopping only at specific places<sup>2</sup>.

Keeping roads out of natural areas is one of the most tractable and cost-effective ways to protect crucial ecosystems. In a world struggling to conserve nature as land-use pressures intensify, managing transport networks is where the rubber meets the road. ■

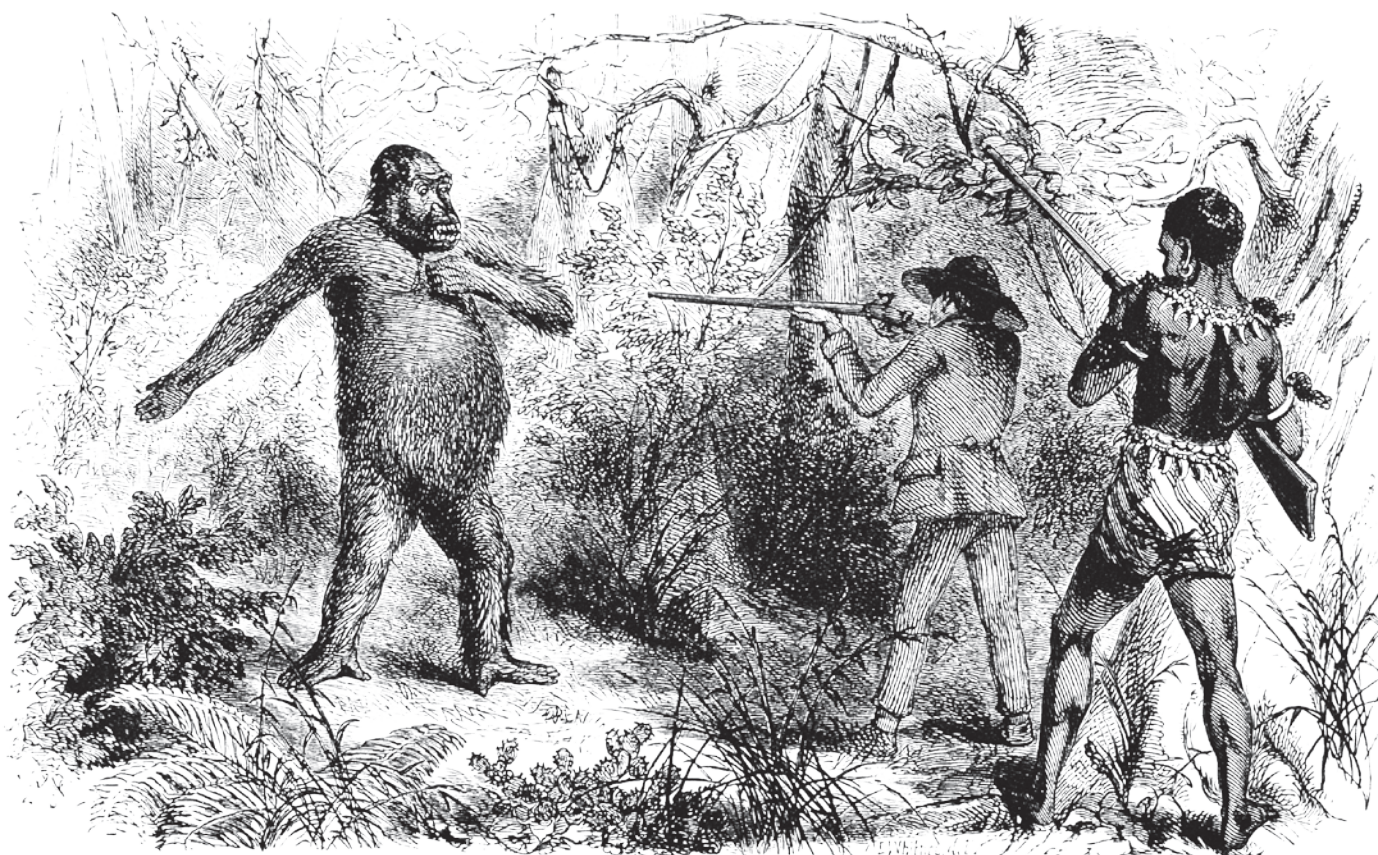
**William F. Laurance** is an Australian Laureate and distinguished research professor of tropical biology at James Cook University in Cairns, Australia. **Andrew Balmford** is professor of conservation science at the University of Cambridge, UK. e-mail: bill.laurance@jcu.edu.au

1. Laurance, W. F. et al. *Science* **291**, 438–439 (2001).
2. Laurance, W. F., Goosem, M. & Laurance, S. G. W. *Trends Ecol. Evol.* **24**, 659–669 (2009).
3. Tilman, D., Balzer, C., Hill, J. & Befort, B. L. *Proc. Natl Acad. Sci. USA* **108**, 20260–20264 (2011).
4. Green, R. E., Cornell, S. J., Scharlemann, J. P. W. & Balmford, A. *Science* **307**, 550–555 (2005).
5. Angelsen, A. & Kaimowitz, D. (eds) *Agricultural Technologies and Tropical Deforestation* (CABI, 2001).
6. Balmford, A., Green, R. & Phalan, B. *Proc. R. Soc. B* **279**, 2714–2724 (2012).
7. Boakes, E. H., Mace, G. M., McGowan, P. J. K. & Fuller, R. A. *Proc. R. Soc. B* **277**, 1081–1085 (2010).
8. Weinhold, D. & Reis, E. *Glob. Environ. Change* **18**, 54–68 (2008).
9. Greenpeace *Partners in Crime: Malaysian Loggers, Timber Markets and the Politics of Self-interest in Papua New Guinea* (Greenpeace International, 2002).



Satellite image of deforestation along the Trans-Amazonian highway in Brazil.





"Mon premier gorille" from adventurer Paul Du Chaillu's colourful 1863 account of his travels in Africa.

## EVOLUTION

# Beastly surprises

Henry Nicholls relishes the history of the 'ape man' who braved the dangers of Gabon — and of Victorian England.

In the middle of the nineteenth century, the famed Victorian publisher John Murray released a title that took Britain by storm. It shot to the top of the best-seller list, dividing the scientific elite and raising troubling questions about the origins of humankind. The book was *Explorations and Adventures in Equatorial Africa* by Paul Du Chaillu.

Never heard of it? That is perhaps due to our obsession with another of Murray's titles, *On the Origin of Species*. It took decades for a true appreciation of Charles Darwin's work to begin to emerge; but *Explorations and Adventures* had a much more immediate and wide-reaching impact. Its enigmatic author became an overnight sensation. Now, journalist Monte Reel goes to the heart of this human drama in *Between Man and Beast*, "the story of a nervy young man who rises, and occasionally falls, in a quest to construct a heroic destiny from scratch".

The first scientific description of the gorilla was given in 1847 to the Boston Natural History Society, but was only read by a handful of the zoological elite. The tribes sharing a

forest with the animal known locally as *njéna* lived in fear of it, believing it lurked in trees, ambushing and choking humans. In 1856, Du Chaillu, the twenty-something son of a French merchant living in West Africa, set out to explore the interior of Gabon — then a cartographic void allegedly rife with cannibals and *njéna*. Three years later, Du Chaillu emerged from the malarious jungle with stories aplenty and the stuffed skins of more than 20 western lowland gorillas. After a brief, largely unsuccessful, attempt to exhibit them in New York, he took his haul to London.

Reel's build-up to this moment is ingenious.



**Between Man and Beast: An Unlikely Explorer, the Evolution Debates, and the African Adventure that Took the Victorian World by Storm**

MONTE REEL  
Doubleday: 2013.  
352 pp. \$26.95

Over the course of several tight, deliciously crafted chapters that flit between Du Chaillu in Africa and the pioneering naturalist Richard Owen in London, Reel sets these two outwardly incongruous characters on a collision course. When Du Chaillu turned up in London, his gorillas were the cargo of Owen's dreams. A sensational exhibition of these dramatically stuffed beasts, he felt, would only support his view (based on differences in brain anatomy) that humans could not be descended from apes. Owen led the young gorilla hunter into the innermost chambers of Britain's scientific establishment, and the heart of Victorian culture itself.

Reel revels in this wider context. He exposes the underbelly of a still-rampant slave trade, takes in a United States on the brink of the civil war, and conjures up the infamous Oxford meeting of the British Association for the Advancement of Science in 1860. There, Thomas Huxley clashed with both Owen and Samuel Wilberforce (the bishop of Oxford) over Darwin's fresh-off-the-press theory. Reel evokes the "soupy fog" of gaslit



London, with its peculiar blend of high society and “smudged underclass” of “ragpickers, costermongers, night-soil men, mud larks, shoeblacks, lamplighters, thimblerriggers”. Whenever Du Chaillu’s compelling drama touches on issues of race and racism, privilege and class, amateurism and professionalism, celebrity and reputation, Reel takes an enriching detour into the details.

With the runaway success of *Explorations and Adventures in Equatorial Africa* and packed lectures at London’s elite scientific institutions, Du Chaillu and his gorillas became an instant cultural phenomenon. They inspired a deluge of (mostly favourable) book reviews, plenty of satirical comment, cartoons, a poem, a sell-out tour of a hirsute ‘freak’ described as “the facsimile of the gorilla” and a popular song and dance known as ‘The Gorilla Quadrille’. They piqued the interest of literary lights such as Charles Dickens, who in 1861 wrote two articles about gorillas in the weekly magazine he edited, *All the Year Round*. In time, Du Chaillu became an inspiration for Arthur Conan Doyle’s action-adventure novel *The Lost World*, Jack London’s wilderness stories, Edgar Rice Burroughs’s tales of Tarzan and Merian C. Cooper’s 1933 film *King Kong*.

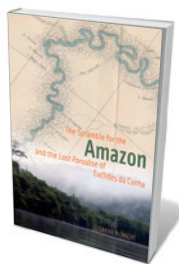
But the sensational colour Du Chaillu used to capture the public imagination began to arouse the suspicion of establishment figures. Perhaps his account was not entirely truthful — or was even completely fabricated. Some wondered whether he might have just bought the skins and skipped the jungle adventure altogether. “He wasn’t just walking the thin line between credibility and bravado; he was dancing on it,” writes Reel. “Every statement of fact in his book was now vulnerable to a contagion of doubt.”

Du Chaillu responded by mounting a second expedition, vowing to collect not only more gorillas but every conceivable snippet of scientific data. Although disastrous, it proved just enough to save his reputation.

Ironically, *Between Man and Beast* itself is not immune from a contagion of doubt. Reel asserts that every scene and quotation “is constructed from historical documents” and “physical descriptions and atmospheric details are rooted in factual evidence”. This is undoubtedly an honest statement, but serious historians will be uncomfortable with Reel’s intense, near-filmic reconstruction of historical events, peppered with detail that is either unverifiable or based solely on Du Chaillu’s own recollections. For everyone else, however, this is what makes the book a supremely entertaining, enlightening and memorable read. ■

**Henry Nicholls** is a freelance journalist based in London. His most recent book is *The Way of the Panda* (Profile Books, 2010). e-mail: [henry@henrynicholls.com](mailto:henry@henrynicholls.com)

## Books in brief



### The Scramble for the Amazon and the *Lost Paradise* of Euclides da Cunha

Susanna Hecht UNIV. CHICAGO PRESS 600 pp. \$45 (2013)

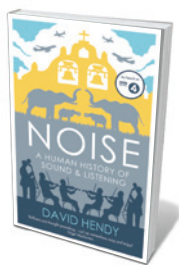
Historian Susanna Hecht charts a brutal nineteenth-century resource grab: the US and European rubber raids in the Amazon. Woven through is the story of great Brazilian environmental writer, geographer and engineer Euclides da Cunha, whose survey expedition down the Amazon River exposed the raids’ grim ecological and human toll. Da Cunha’s account, *Lost Paradise*, masterfully mixes biology, geography and philosophy, but remains unfinished: he died violently. A journey into South America’s heart of darkness.



### Turned Out Nice Again: On Living With the Weather

Richard Mabey PROFILE BOOKS 160 pp. £8.99 (2013)

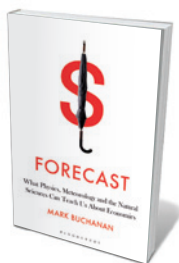
As denizens of an island prone to flooding, gales, drizzle and the occasional halcyon day, Britons obsess about meteorological vagaries. Nature writer Richard Mabey celebrates this preoccupation with the weather that, glue-like, bonds a nation. In a lyrical 90 pages, he takes us from freak events such as the January 1940 ice storm that saw cats “iced to branches”, to retinal detachment during extreme low fronts and the soggy impact of climate change. Mabey’s veerings — from forecaster-shamans to naturalist Gilbert White’s frozen pisspot — are as gloriously mercurial as the British weather.



### Noise: A Human History of Sound and Listening

David Hendy PROFILE BOOKS 402 pp. £16.99 (2013)

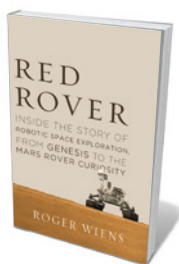
An acoustic history spanning some 100,000 years, this BBC Radio 4 tie-in is a smorgasbord of sound. David Hendy leads us into the ‘voiceprints’ of human prehistory through the music of wind-torn forests, calling animals and stridulating insects that must have inspired our early ancestors. He then tackles the roots of oratory; the aural assault course that was ancient Rome; the specialized ‘smart’ acoustics of churches and mosques; noise anxiety in an urbanizing, militarized world; the rise of recording technology; and the contemporary search for silence in the soundscape.



### Forecast: What Physics, Meteorology, and the Natural Sciences Can Teach Us About Economics

Mark Buchanan BLOOMSBURY 272 pp. £18.99 (2013)

Physicist Mark Buchanan rethinks the wobbly phenomenon of international finance. Markets, he argues, do not self-regulate. They are as vulnerable to severe ‘weather events’ as Kansas: prone, like any system, to positive feedbacks in which small variations can lead to sudden change. Buchanan argues for a physics-flavoured disequilibrium approach that would improve forecasts through the creation of models that are fine-tuned to recognized shifts in trade, speculation and other financial activities.



### Red Rover: Inside the Story of Robotic Space Exploration, from Genesis to the Mars Rover Curiosity

Roger Wiens BASIC BOOKS 256 pp. \$25.99 (2013)

Space hardware is in boom mode, points out geochemist — and principal investigator for Curiosity’s ChemCam instrument — Roger Wiens. Meshing the blow-by-blow science with memoir, his account begins with childhood rocket modelling and telescope building, moves on to the 2001 Genesis mission and culminates with the financial and technological rollercoaster in the lead-up to Curiosity. A tribute to human ingenuity in NASA’s “faster, better, cheaper” era.





A baby bonobo at the Lola Ya Bonobo sanctuary in the Democratic Republic of the Congo.

## PRIMATOLOGY

# A wild empathy

Christopher Boehm relishes a wide-ranging assessment of primate morality.

Frans de Waal's latest book, *The Bonobo and the Atheist*, is both an exceptionally good read and a tour de force of scholarship. In it, de Waal states his argument for the evolution of human empathy with the sophistication of a well-grounded, risk-taking scientist who can venture into philosophy.

De Waal draws on his own ethological research with chimpanzees and bonobos, as well as on biology and evolutionary psychology, to probe empathy as a key precursor for moral behaviour. The origins of human morality are not new territory for the primatologist and ethologist, who engaged with the question in *Peacemaking Among Primates* (Harvard University Press, 1990) and who, with Jessica Flack, has examined the evolutionary building blocks of morality. The novel element in *The Bonobo and the Atheist* is that de Waal analyses today's moral landscape, in particular the schism between militant atheism and religion.

De Waal views extreme strains of atheism as getting "all worked up about the absence of something", at one point using the fanciful device of a talking bonobo as his mouthpiece (hence the book's title). His view is that religion is undeniably in our bones — even

➔ **NATURE.COM**  
For Christian  
Keyesers on Frans  
de Waal, see:  
[go.nature.com/plb4cq](http://go.nature.com/plb4cq)

though evidence of primate precursors seems less than substantial. This does not mean that he is pro-religion, however. *The Bonobo and the Atheist* is permeated with the ethos of secular humanism, using the Renaissance painter Hieronymus Bosch's *The Garden of Earthly Delights* — a vision of humanity freed from narrow moral constraints — as a touchstone for his arguments.

In his discussion of empathy and morality, de Waal has little time for what he calls "veneer theories" that reduce altruism to 'natural' selfishness. As he shows, human altruism has analogues in a wide range of species, even though sterile ants' care for the offspring of their queen can hardly be labelled empathy. When dolphins assist humans struggling in the water, we may at least suggest some basic similarities. But when a chimpanzee, sharing more than 95% of our DNA, helps an unrelated member of its group to lick a wound it cannot reach, a type of empathy very near the



**The Bonobo and the Atheist: In Search of Humanism among the Primates**  
FRANS DE WAAL  
W. W. Norton: 2013.  
320 pp. \$27.95,  
£18.99

human is surely coming into play.

Many evolutionists favour chimpanzees as ancestral models. Whereas de Waal does look frequently to chimpanzees as exemplars of primate altruism, he champions the less violent bonobo — not least because its habitat, like that of our common primate ancestor, remains the tropical forest, whereas chimpanzees and humans have evolved into ecological generalists.

De Waal looks to mothering and infant care by non-kin, a basic form of empathy discussed by primatologist Sarah Hrdy in *Mothers and Others* (Harvard University Press, 2011), as the foundation of human altruism and complex cooperation, and as his prime evolutionary building block for morality. He also emphasizes the importance of emotion in moral choices, citing the work of psychologist Jonathan Haidt, author of *The Righteous Mind* (Allen Lane, 2012). Haidt's empirical investigations of subjects' disgusted reactions to incest demonstrate that when it comes to morality, raw emotions trump rationality.

*The Bonobo and the Atheist* does leave a gap of sorts in explaining moral origins. In discussing his building blocks, de Waal seems relatively unconcerned with the actual mechanisms of natural selection that keep them in place, be this through kin selection, mutualism, reciprocal altruism, group selection or social selection.

These theories can help to explain why nice guys don't necessarily finish last. Group selection theory may be making a comeback, as seen in a 2010 paper in *Nature* by E. O. Wilson and others; there are also social-selection arguments in Richard Alexander's *The Biology of Moral Systems* (Aldine Transaction, 1987), where he suggests that people's altruistic reputations may seriously enhance reproductive success. In my own *Moral Origins* (Basic Books, 2013) I suggest that social selection in the form of groups punishing moral deviants has profoundly shaped human gene pools, and given us a conscience — an argument de Waal does discuss.

De Waal's style is refined yet lively; his humanism both highbrow and lowbrow. One minute he is referencing the art of Bosch, the next comparing himself to an Australian toilet frog, showered over the decades by distastefully superficial sociobiological explanations of human altruism. *The Bonobo and the Atheist* is a synthesis on all levels, masterfully marshalling ethology, psychology, philosophy and anthropology in its drive to understand ourselves through the lens of other primates. ■

**Christopher Boehm** is professor of biological sciences and anthropology and director of the Goodall Research Center at the University of Southern California, Los Angeles, USA.  
e-mail: [cboehm1@msn.com](mailto:cboehm1@msn.com)

## EVOLUTIONARY BIOLOGY

# Gentle giant of genetics

**Oren Harman** assesses the first biography of biologist W. D. Hamilton, the 'greatest Darwinian since Darwin'.

He was socially awkward, charmingly accident-prone, honest, modest, a champion of underdogs. A romantic who saw the human condition as essentially tragic, he felt more at home in the Amazon bush than with people, judging culture "a braggart"; the genes, to his mind, really ran the show. And his theory of inclusive fitness, many hold, is the single most original development in evolutionary theory since Darwin, explaining the workings of natural selection through the 'gene's eye' point of view.

The British evolutionary biologist W. D. "Bill" Hamilton (1936–2000) was a naturalist of legendary knowledge, especially of insects, and a complex man, both rugged and compassionate. An uncompromising truth-seeker, he set the highest of standards for his scientific papers, which were often rejected, famously by *Nature*. Now, in *Nature's Oracle*, sociologist Ullica Segerstrale presents the first biography of this gentle giant with a poet's soul.

Hamilton asked big questions and provided answers just as dramatic. Where does altruism come from? Why does sex exist? Why are females attracted to males? How does cooperation in nature arise? Marshalling mathematics, and later computer modelling, in the service of an acute power of observation and an ability to see the world from the point of view of the tiniest biological creatures and entities — the wasp, the virus, the gene — Hamilton reached often counter-intuitive solutions. Sex is merely an adaptation to avoid bugs, with ourselves and other diploids as "guilds of genotypes", as he put it, "committed to free, fair exchange of biochemical technology for parasite exclusion". Cooperation is a game played by the self-interested, nothing more; aesthetics is a signal for good genetics, meant to garner mates.

An intellectual and physical daredevil, Hamilton was also often distracted and sometimes depressed. Sporting a caveman-like shock of thick white hair, he was famous for mumbling inaudibly, his back to the audience, during virtually incomprehensible lectures in the United States and Britain. He felt alone and uncomprehended. Despite his unusual generosity in giving credit to others for his own fountain-like ideas, Segerstrale also uncovers an exaggerated jealousy for primacy, and an 'outsider' mentality that persisted long after he had become celebrated.

It is not easy to remain unmoved by

Hamilton's blend of obsession, sincerity, vulnerability and loneliness. Even more difficult is the task of writing a biography of a man unusually gifted as a writer himself, who left behind a brutally honest intellectual autobiography, *Narrow Roads of Gene Land*, in two volumes (a third followed, edited posthumously). Readers will be excused for noting that Segerstrale's biography seems to rely rather heavily on these. "I tend to be more like a bird than alligator — hatching my ideas to maturity before letting them loose," Hamilton wrote of himself. Passing by a yellow ragwort struggling to grow through a crack in London's asphalt, he noted, "a screaming child would not have touched my heart in the same way". It's a tough charge to improve on that.

In places, under-reliance on the secondary literature leads to either conjecture (that J. B. S. Haldane sought to prove altruism compatible with natural selection because of his left-leaning political tendencies) or factual fuzziness (missing the crucial impact of George Price's covariance paper on Hamilton's view on group selection). A somewhat incomplete treatment of Hamilton's intellectual legacy also leaves the reader lacking the



**Nature's Oracle: A Life of W. D. Hamilton**  
ULLICA SEGERSTRALE  
Oxford University Press: 2013. 441 pp.  
£25, \$35



W. D. Hamilton a few years before his death.

wider context. In the past few years, the pages of *Nature* have seen a hard-fought debate between proponents of Hamiltonian kin selection and its challengers, but Segerstrale chooses not to delve deeply into the controversy. Had she done so, she would have been able to show the extent to which E. O. Wilson's current views on the evolution of eusociality represent a return to Hamilton's ideas.

Hamilton was indeed an inveterate seeker of patterns, a man with "evolution on the brain", as Segerstrale puts it, but he sometimes took this too far. His application of extreme geneticism to humans is an example. Wary of the adverse effects of modern medicine on the gene pool, he called unapologetically for infanticide of the genetically weak. This is one instance in which the otherwise singular Hamilton is anything but unique, joining a long line of naive champions of eugenics — and it demands more detached, hard-nosed criticism than offered here.

Still, combing through letters and scrupulously interviewing family members, colleagues and friends, Segerstrale succeeds in bringing Bill Hamilton, warts and all, to what will be an appreciative audience. "I am ashamed that in practice I find myself much more selfish than my ideal," the man who tackled altruism confessed touchingly in a long, paranoid letter to John Maynard Smith in 1977, accusing Smith of not giving him sufficient credit for the idea of inclusive fitness. Hamilton's mother, a medical doctor, saw signs of Asperger's syndrome (an autism spectrum disorder) in characteristics such as his obsession with priority, preoccupation with patterns, minuscule handwriting and focus on number and detail. He himself thought, tellingly, that being a 'thing', rather than a 'people', person accounted for his insight into nature.

Whatever the truth in these suppositions, Segerstrale's biography of Hamilton is ultimately a celebration of the life in science of a deeply feeling human being — an "evolution sufferer" as he called himself. In the grand scheme of nature, Hamilton knew, a human life is desperately transient. So did his favourite poet, A. E. Houseman:

*For nature, heartless, witless nature,  
Will neither care nor know  
What stranger's feet may find the meadow  
And trespass there and go,  
Nor ask amid the dews of morning  
If they are mine or no.*

Despite it all, Bill Hamilton's life is one to remember. ■

**Oren Harman** is professor of the history of science at Bar-Ilan University, Israel, and author of a biography of George Price, *The Price of Altruism* (W. W. Norton, 2010). [oren.harman@gmail.com](mailto:oren.harman@gmail.com)



# Correspondence

## Medical supplies in Iran hit by sanctions

The delivery and manufacture of vital medical supplies in Iran is being impeded by bilateral sanctions on the country's economy, imposed to force Iran to comply with international rules on nuclear programmes. This situation is causing a humanitarian disaster and contravenes the Universal Declaration of Human Rights.

Medicines are not themselves subject to sanctions, but the licensing, purchase and shipment of imported medicines have all been hit. This, combined with the potential threat of sanctions by international pharmaceutical companies and banks, has led to a shortage of specific drugs and medical facilities in Iran over the past few months, increasing drug prices by 50% and affecting more than 6 million chronically ill patients.

Experience in Iraq, Cuba, Libya and the former Yugoslavia indicates that sanctions seldom meet their political objectives. But they have caused large-scale humanitarian disasters in those countries (see, for example, M. M. Ali and I. H. Shah *Lancet* 355, 1851–1857; 2000).

The Iranian authorities must take immediate action to block the effect of sanctions on the supply of medicines and of materials for hospitals and laboratories.

The international community needs to establish criteria for guaranteeing that medical products will always be exempt from sanctions.

**Ali Gorji** *Epilepsy Research Centre, University of Münster, Münster, Germany.*  
gorjial@uni-muenster.de

## Fisheries: broaden the arguments

The split over whether catch is a useful indicator of fish abundance (D. Pauly, R. Hilborn and T. A. Branch *Nature* 494, 303–306;

2013) is a false dichotomy. Not because a realistic position lies somewhere in between, but because the focus is too narrow.

In the whole-ecosystem approach to fisheries management, data on all species, including by-catch species, must be considered. A broad range of information is therefore used to assess the health of fisheries worldwide, from no-data situations that rely on precautionary and risk-based approaches, to data-poor fisheries in which catch data are unreliable, as well as data-rich fisheries that are well supported by catch data and fishery-independent surveys.

As climate change alters the distribution and abundance of different species, catch data may become less useful and new methods will be needed for assessing fish stocks. However, even the best estimates of fish numbers will not improve fisheries' health without good leadership, robust social capital, and effective conservation and stock-rebuilding policies.

**Alistair J. Hobday** *CSIRO*

*Marine and Atmospheric Research, Hobart, Tasmania, Australia.*

alistair.hobday@csiro.au

**Rodrigo H. Bustamante, Éva**

**E. Plagányi** *CSIRO Marine and Atmospheric Research, Brisbane, Queensland, Australia.*

## Fisheries: manage declines

Catch data are first and foremost measures of fishery output (D. Pauly, R. Hilborn and T. A. Branch *Nature* 494, 303–306; 2013). A long-term decline in catch can be a warning that conditions have changed, either in the abundance and productivity of fish populations or in fishing practices. Although increases in fishing effort can sometimes temporarily restore yields, it is important to investigate the causes of declines.

When time series of commercial catch data are the only information available, there are ways to supplement them with life-history information (such as maximum body length) to expose important changes in population abundance (see C. Costello *et al. Science* 338, 517–520; 2012).

In areas of the world that cannot afford investigations of abundance, a reduction in fishing intensity is prudent precautionary management (see *Code of Conduct for Responsible Fisheries* UN Food and Agriculture Organization; 1995).

**Brian R. MacKenzie, Mark**

**R. Payne** *National Institute for Aquatic Resources, Technical University of Denmark, Charlottenlund, Denmark.*  
brm@aqu.dtu.dk

## Australia's grant system wastes time

We found that scientists in Australia spent more than five centuries' worth of time preparing research-grant proposals for consideration by the largest funding scheme of 2012. Because just 20.5% of these applications were successful, the equivalent of some four centuries of effort returned no immediate benefit to researchers and wasted valuable research time. The system needs reforming and alternative funding processes should be investigated.

We surveyed a representative sample of Australian researchers and found that preparing new proposals for the National Health and Medical Research Council's project grants took an average of 38 working days; resubmitted ones took 28 days on average. Extrapolating this to all 3,727 submitted proposals gives an estimated 550 working years of researchers' time (95% confidence interval, 513–589), equivalent to a combined annual salary cost of Aus\$66 million (US\$68 million). This exceeds the total salary bill (Aus\$61.6 million) at

Melbourne's Walter and Eliza Hall Institute, a major medical-research centre that produced 284 publications last year.

The grant proposals we analysed were typically 80–120 pages long. If these were more focused, it would reduce preparation costs and could improve the quality of peer review by reducing workloads.

**Danielle L. Herbert, Adrian G. Barnett, Nicholas Graves** *Queensland University of Technology, Brisbane, Australia.*  
n.graves@qut.edu.au

## Outdated taxonomy blocks conservation

We agree with Frank Zachos and colleagues that splitting mammalian species taxonomically without scientific justification could impede their conservation (*Nature* 494, 35; 2013). But so does uncritically lumping them together.

Mammalogy is beleaguered by a dogmatic regard for mid-twentieth-century propositions, which were seldom based on critical study and lacked phylogenetic information. Species were lumped together and incorporated into influential classification checklists to simplify regional faunas and make them more manageable for non-taxonomists.

Modern integrative approaches have shown that this tactic has hidden an incommensurable number of distinctive species from conservation efforts (W. R. Morrison *et al. Biol. Conserv.* 142, 3201–3206; 2009), thereby increasing the risk of extinctions.

Scepticism should be accorded to any taxonomy that is not based on comprehensive revisionary work, phylogenetic studies or, ideally, both.

**Eliécer E. Gutiérrez, Kristofer M. Helgen** *National Museum of Natural History, Smithsonian Institution, Washington DC, USA.*  
gutierrez@si.edu

## 3D without the glasses

A glasses-free three-dimensional display technology has been invented that may be an exciting alternative to current solutions for mobile devices. It makes use of an optical effect known to school physics students. [SEE LETTER P.348](#)

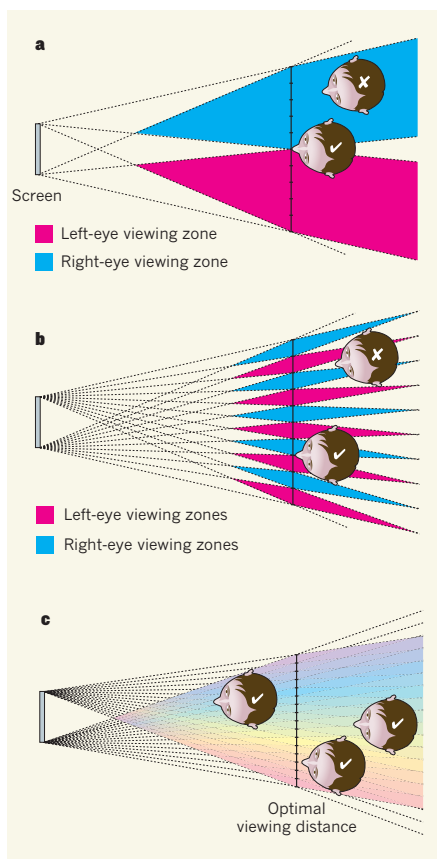
NEIL A. DODGSON

We are all familiar with the glasses needed to watch three-dimensional (3D) movies. But glasses are inappropriate for many possible uses of a 3D display. There are a range of technologies that can achieve 3D without the glasses. On page 348 of this issue, Fattal *et al.*<sup>1</sup> describe a new form of illumination device that could help to achieve it for mobile devices.

Humans see the world stereoscopically. That is, our two eyes see two slightly different views of the world because they are separated by about 63 millimetres<sup>2</sup>. Our brains combine these two views into an internal 3D model of whatever we are seeing. Conventional 2D displays provide only a single flat image, so the two eyes both see the same picture on the screen. A 3D display device must instead present a different image to each eye.

Glasses-based 3D systems can work in several ways. One way, known as active shuttering, uses rapidly flickering shutters in the glasses, synchronizing them with rapid alternation of left and right images on the screen to give a different image to each eye. This is how most home 3D televisions work. A cheaper solution is passive polarization<sup>3</sup>, in which the two lenses in the glasses are polarized in different directions, with the display showing two images simultaneously, one with each polarization. This is how most 3D cinema has worked since the 1950s. And then there is the cheap and cheerful anaglyph solution: the familiar red-green glasses<sup>4</sup>, where each eye's image is produced using filters of different colours (red and green). Anaglyph was used for the earliest experiments in 3D movies, in 1922, but was quickly found to cause headaches, hence the migration to polarized systems. However, anaglyph is still used as a cheap alternative, especially in print media.

Consider now how you could achieve 3D without using any special glasses. You would need to send a different image to each eye, but the mechanism for doing this now needs to be in the display itself. The individual pixels in the display therefore need directional control. Conventional 2D displays have pixels that send light in all directions. A 3D display needs pixels that send light in carefully constrained



**Figure 1 | Autostereoscopic effects.** a–c, Three examples of a 3D imaging display in which the 3D effect is encoded in the display and can be viewed with no need for glasses. **a**, Only two viewing zones, one for each eye. A viewer with their nose in the right place sees 3D. Anyone else sees just 2D: the same image to each eye. **b**, Only two views, but the zones repeat. A viewer in the right place sees 3D. A viewer in the wrong place sees pseudoscopically: the wrong image to each eye. **c**, Sixteen views. Everyone sees a 3D image. This is the effect of multiview autostereoscopic displays such as that proposed by Fattal and colleagues<sup>1</sup>.

directions, so that different light reaches each eye and the viewer's two eyes thus see different images on the screen. You might assume that just two constrained directions are needed: one for each eye (Fig. 1a,b). But this is not the case, because the display does not know where your eyes are and therefore it has to provide imagery across the entire viewing space, to give

the correct 3D effect no matter where the eyes are placed.

The ideal solution is to have pixels that are able to display different colours in every possible direction. This can be achieved using holographic techniques but, currently, holography is practical only for still images. For updatable displays, or 3D video, we need the next best thing to holography: autostereoscopic multi-view 3D display. That is, a stereoscopic display (a different picture to each eye) that works automatically (without requiring the user to wear any special device) and displays multiple different images on one display screen, each visible from particular places in front of the screen. Figure 1c shows such a display with 16 different viewing zones<sup>5</sup>. In each zone, a different image is visible on the screen. Every viewer sees in 3D (a different picture to each eye) and each viewer sees 3D from their particular position (different pictures as you move your head from side to side).

There are several ways to achieve this for large-scale displays: optical elements such as lenticular arrays<sup>6</sup> or parallax barriers<sup>7</sup> in front of a liquid-crystal display panel, multiple projectors behind a large lens system<sup>8</sup>, or rapid display of multiple images combined with complex optical switching<sup>9</sup>. Lenticular arrays are a relatively successful mechanism for desktop 3D. They have the advantage that their constituent lenticules (tiny abutting cylindrical lenses) comprise a thin sheet in front of a liquid-crystal display, making them cheap to manufacture. The disadvantage of lenticular arrays, shared with parallax barriers, is that they reduce the native resolution of the underlying display by a factor equal to the number of distinct viewing directions.

The most familiar glasses-free mobile display is the Nintendo 3DS, which uses a parallax-barrier system to create two viewing zones. With only two views, the user has to be roughly the right distance from the screen and must ensure that their nose is on the boundary between the two zones, so that each eye sees the appropriate view. The optics of these devices mean that the viewing zones are repeated as the head is moved left–right (Fig. 1b). Lenticular arrays, producing the same kind of effect, have also been used for other mobile devices.

Fattal *et al.* offer a different approach to



producing an autostereoscopic multiview display. They use a light guide combined with diffraction gratings (see Fig. 1a,b of the paper<sup>1</sup>) to direct light in specific directions. Their demonstration models can send light in 14 distinct viewing directions, and they believe that the design can be made to produce up to 64. The implication is that their devices could be used to produce an autostereoscopic display with up to 64 distinct viewing zones. With that many zones, the inter-zone spacing could be considerably smaller than the distance between the eyes, producing a smooth 3D effect that would be similar to viewing a white-light hologram.

However, there are considerable challenges to overcome before we will see hologram-like displays on a mobile device. First, we need to make sure that 3D will not compromise spatial resolution. Most current mobile devices have pixels about 250 micrometres across. The new illumination system has individual direction

pixels 36 micrometres across, theoretically allowing us to pack up to 48 directions into the space of a single pixel. This device thus seems to give us the desired directional control with no loss of spatial resolution. Next, the 3D effect must not compromise image quality. The example images in the paper show that there is considerable work to be done to improve quality to an acceptable level. Then, the display must have the 64 images to display. This is straightforward, although computationally expensive, with computer-generated imagery but is a significant challenge for live action.

Finally, the new illumination system must be manufactured reliably, robustly and in quantity. This is a matter of careful engineering that can take a long time. I am reminded of the nine-year development of Texas Instruments' digital micro-mirror device, now widely used in digital projectors, between the invention in 1987 and the first commercial product in 1996.

By analogy, if the authors can solve the practical problems, then they have a compelling alternative to existing 3D display technology. All that remains is the more nebulous question of whether humans want or need 3D displays. Time will tell. ■

**Neil A. Dodgson** is at the Computer Laboratory, University of Cambridge, Cambridge CB3 0FD, UK.  
e-mail: neil.dodgson@cl.cam.ac.uk

1. Fattal, D. *Nature* **495**, 348–351 (2013).
2. Dodgson, N. A. *Proc. SPIE* **5291**, 36–46 (2004).
3. Land, E. H. US patent 2,099,694 (1937).
4. du Hauron, L. D. US patent 544,666 (1895).
5. Dodgson, N. A. *Appl. Opt.* **35**, 1705–1710 (1996).
6. van Berkel, C. & Clarke, J. A. *Proc. SPIE* **3012**, 179–186 (1997).
7. Konrad, J. & Halle, M. *IEEE Signal Process. Mag.* **24**(6), 97–111 (2007).
8. Kawakita, M. et al. *Proc. SPIE* **8288**, 82880B (2012).
9. Travis, A. R. L. *Appl. Opt.* **29**, 4341–4343 (1990).

## STEM CELLS

# Painkillers caught in blood-cell trafficking

**Haematopoietic stem and progenitor cells move from the bone marrow into the circulation to replenish normal blood-cell levels. Inhibiting a prostaglandin-mediated signalling pathway may promote this process. [SEE LETTER P.365](#)**

JASON M. BUTLER & SHAHIN RAFII

Stem and progenitor cells of the blood system reside in the safe haven of their immediate surroundings — a micro-environment that contains a complex network of 'niche' cells. Interaction with niche cells prevents unauthorized egress of these haematopoietic stem and progenitor cells into the circulation, whereas their mature counterparts are readily launched there. How the various niche cells decide to relocate some cells but not others, to meet the demand for blood cells under normal and stressful conditions, is not fully understood. On page 365 of this issue, Hoggatt *et al.*<sup>1</sup> provide one notable clue\*.

The authors show that interfering with the interaction between the messenger lipid molecule prostaglandin E<sub>2</sub> (PGE<sub>2</sub>) and its G-protein-coupled receptor EP4 stimulates the mobilization of haematopoietic stem cells and haematopoietic progenitor cells (HSCs/HPCs) from the bone marrow to the circulation. This effect was evident following deletion of the *EP4* gene in mice; using EP4 antagonists; and after lowering PGE<sub>2</sub> production by treating mice

with non-steroidal anti-inflammatory drugs (NSAIDs), including aspirin, ibuprofen and meloxicam.

Individually, NSAIDs and EP4 antagonists were marginally effective in forcing HSC/HPC egress, but their effect was considerably improved when they were used together with other mobilizing agents — G-CSF and a CXCR4 antagonist. Most importantly, the authors report such enhanced effects of NSAIDs, alone or in combination with G-CSF in non-human primates and human subjects. This sets the stage for clinical studies to interrogate the potential of PGE<sub>2</sub> for optimizing HSC/HPC mobilization protocols as part of therapies involving myeloablation, which involves the destruction of diseased blood cells followed by induced production of new blood cells from the stem-cell pool.

The mechanism by which PGE<sub>2</sub>–EP4 signalling regulates blood-cell trafficking is complex and involves balancing HSC/HPC mobilization both directly, through cell-specific signalling, and indirectly, by activating stromal cells of the niche. Together with earlier work<sup>2,3</sup> showing that, by contrast, short-term exposure of haematopoietic cells to PGE<sub>2</sub> increases their migration and engraftment, these

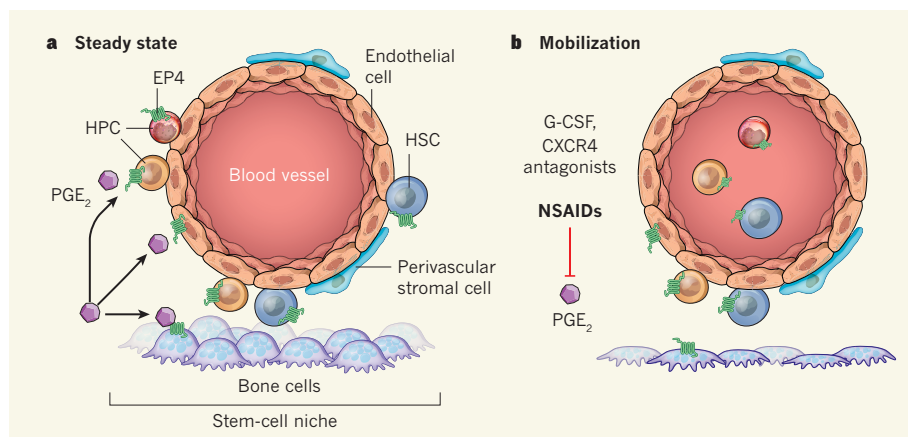
findings qualify the PGE<sub>2</sub>–EP4 signalling hub as a gatekeeper of the haematopoietic microenvironment (Fig. 1).

Uncovering the molecular pathways involved in the efficient mobilization and engraftment of HSCs/HPCs has major implications for the treatment of various cancers with myeloablative therapies<sup>4,5</sup>. High blood levels of a range of factors — including G-CSF, CXCR4 antagonists, numerous CC and CXC chemokines<sup>4</sup>, VEGF-A and PlGF (ref. 5) — were known to stimulate the mobilization of HSCs/HPCs localized to specific haematopoietic niches. However, the identity of factors that could modulate both mobilization and engraftment of these cells were not known.

PGE<sub>2</sub> is one such factor, and, by interacting with EP4, it regulates bidirectional trafficking of HSCs/HPCs. Indeed, the effects of PGE<sub>2</sub> could be biphasic. Brief stimulation of HSCs/HPCs with this molecule augments engraftment and reconstitution of blood cells by accelerating haematopoiesis<sup>6</sup>. This function of PGE<sub>2</sub> is mediated by activation of the Wnt signalling pathway to control proliferation and programmed death of HSCs/HPCs<sup>7</sup>. Non-stop stimulation, however, may inhibit HPC expansion and recovery of haematopoiesis<sup>8</sup>, as well as prevent the trafficking of HSCs/HPCs, partly by CXCR4 downregulation. Therefore, capitalizing on the potential of the PGE<sub>2</sub>–EP4 signalling pathway to augment haematopoiesis may require a two-pronged approach in which EP4 is initially activated to trigger migration and engraftment of HSCs/HPCs, and then inhibited to encourage HPC expansion and haematopoietic reconstitution.

An intriguing question is exactly how NSAIDs modulate trafficking of the HSCs/HPCs within their microenvironment. EP4 is expressed by bone cells called osteoclasts and osteoblasts; by endothelial cells, which line the

\*This article and the paper under discussion<sup>1</sup> were published online on 13 March 2013.



**Figure 1 | NSAIDs are gatekeepers of the haematopoietic microenvironment.** **a**, Prostaglandin E<sub>2</sub> (PGE<sub>2</sub>) can bind to its receptor EP4 on the surface of haematopoietic stem and progenitor cells (HSCs/HPCs), as well as on the surface of niche cells within the HSC/HPC microenvironment, which include perivascular stromal cells, endothelial cells that line blood vessels and cells of the bone lineage. The signalling cascade initiated by PGE<sub>2</sub>–EP4 interaction inhibits mobilization of HSCs/HPCs into the circulation. **b**, Hoggatt *et al.*<sup>1</sup> show that NSAIDs, by inhibiting PGE<sub>2</sub>–EP4 interactions and in part by thinning of the bone cells, can synergize with the mobilizing agents G-CSF and CXCR4 antagonists. This promotes egress of HSCs/HPCs into the circulation by weakening their tethering to the niche cells.

blood vessels; by subsets of haematopoietic cells; and most probably by perivascular cells, which occur in peripheral blood vessels. Hoggatt *et al.* show that treatment with NSAIDs results in thinning and loss of the cells of the bone lineage, diminishing their capacity to retain HSCs/HPCs within the bone marrow. Moreover, they find that NSAIDs decrease bone-derived osteopontin, a molecule that regulates mobilization of haematopoietic cells. These painkillers might therefore primarily target bone cells to decrease the tethering of HSCs to stromal cells.

Nonetheless, PGE<sub>2</sub> might also induce mobilization by directly affecting perivascular and endothelial cells. Notably, sinusoidal endothelial cells line small blood vessels called sinusoids. The sinusoids, which form branching structures in much of the bone-marrow vasculature, are not mere passive conduits for delivering oxygen and nutrients, but rather represent specialized blood vessels that form an instructive 'vascular niche' in the haematopoietic organs. Normally, activated endothelial cells produce defined blood-related growth factors (angiocrine factors) to balance the self-renewal and differentiation of HSCs<sup>9–16</sup>. One angiocrine factor is PGE<sub>2</sub>, through which endothelial cells could modulate HSC/HPC trafficking. Indeed, during recovery from myeloablative therapy, the induction of angiocrine factors — including PGE<sub>2</sub> — in endothelial cells regulates trafficking and regeneration of HSC/HPCs<sup>11–14</sup>. Unravelling the mechanism by which PGE<sub>2</sub> polices the trafficking of HSCs/HPCs within the complex network of intertwined niche cells will require the deletion of EP4 in specific niche cells, including endothelial cells, perivascular cells and cells of the bone lineage.

Notwithstanding the potential benefits of NSAIDs, these drugs may act as a

double-edged sword during haematopoietic recovery. Because activation of PGE<sub>2</sub>–EP4 signalling promotes blood-vessel formation, NSAIDs might interfere with the regeneration of endothelial cells, thus impairing haematopoietic reconstitution. These complex issues could also be addressed by differentially deleting EP4 in various bone-marrow stromal niche cells, specifically endothelial cells during recovery from myeloablation. In addition, selective deletion of prostaglandin-E synthase enzymes in various niche cells will define the specific niche cells that functionally deploy PGE<sub>2</sub> to activate HSCs/HPCs.

The present findings are relevant to whether NSAIDs can restore the mobilization of

HSCs/HPCs in patients who are resistant to treatment with G-CSF and CXCR4 antagonists. It is plausible that resistance to these factors is not due to intrinsic HSC/HPC defects, but rather reflects impaired instructive functions of haematopoietic niche cells. Thus, NSAIDs could pave the way for defining alternative routes to safely mobilize HSCs/HPCs through manipulation of haematopoietic niche cells. These observations also shed light on the mechanism by which NSAIDs may contribute to protection against cardiovascular disorders, by mobilizing haemangiocytes — HPCs that promote organ-specific formation of functional blood vessels<sup>17</sup>. ■

**Jason M. Butler and Shahin Rafii** are in Weill Medical College, Cornell University, and the Howard Hughes Medical Institute, New York, New York 10065, USA. e-mail: srafiu@med.cornell.edu

- Hoggatt, J. *et al.* *Nature* **495**, 365–369 (2013).
- Hoggatt, J., Singh, P., Sampath, J. & Pelus, L. M. *Blood* **113**, 5444–5455 (2009).
- North, T. E. *et al.* *Nature* **447**, 1007–1011 (2007).
- Bonig, H. & Papayannopoulou, T. *Leukemia* **27**, 24–31 (2013).
- Hattori, K. *et al.* *Nature Med.* **8**, 841–849 (2002).
- Goessling, W. *et al.* *Cell Stem Cell* **8**, 445–458 (2011).
- Goessling, W. *et al.* *Cell* **136**, 1136–1147 (2009).
- Hoggatt, J. & Pelus, L. M. *Leukemia* **24**, 1993–2002 (2010).
- Butler, J. M. *et al.* *Cell Stem Cell* **6**, 251–264 (2010).
- Hooper, A. T. *et al.* *Cell Stem Cell* **4**, 263–274 (2009).
- Kobayashi, H. *et al.* *Nature Cell Biol.* **12**, 1046–1056 (2010).
- Butler, J. M. *et al.* *Blood* **120**, 1344–1347 (2012).
- Heissig, B. *et al.* *Cell* **109**, 625–637 (2002).
- Kimura, Y. *et al.* *PLoS ONE* **6**, e26918 (2011).
- Ding, L. *et al.* *Nature* **481**, 457–462 (2012).
- Ding, L. *et al.* *Nature* <http://dx.doi.org/10.1038/nature11885> (2013).
- Jin, D. K. *et al.* *Nature Med.* **12**, 557–567 (2006).

## EARTH SCIENCE

## Mainly in the plain

**The finding that global mass loss from landscapes is dominated by physical erosion and chemical weathering from flat terrain, rather than from mountains, challenges our understanding of how Earth's surface evolves.**

**JAMES W. KIRCHNER & KEN L. FERRIER**

Inexorably, everything falls apart. As individuals, we humans last only some decades or so. Our most Ozymandian creations crumble into oblivion over hundreds or thousands of years. And on million-year timescales, even the landscapes on which we build our civilizations will disappear, gradually consumed by physical erosion and chemical dissolution. How quickly these processes transform Earth's surface has been a question of great interest to geologists and geochemists.

But they may have mostly been looking for the answer in the wrong places. Willenbring *et al.*<sup>1</sup> report in *Geology*. Geologists love mountains, and have focused their attention on steep terrain, where rapid erosion is thought to drive rapid chemical weathering<sup>2</sup>. However, Willenbring *et al.* say that, on a global basis, there is just not enough mountainous terrain to account for much of the global flux of sediment (from erosion) or solutes (from chemical weathering) reaching the oceans.

This argument is likely to arouse interest, and some measure of controversy, because



it concerns not only the evolution of Earth's surface, but also its long-term climate. When calcium and magnesium are dissolved from silicate minerals and later deposited as carbonates on the ocean floor, carbon dioxide is removed from the ocean–atmosphere system. This mechanism may stabilize global climate over the long term if the climatic shifts that accompany higher atmospheric CO<sub>2</sub> levels also accelerate global erosion and weathering fluxes.

Willenbring and colleagues' analysis is based on a global compilation of what geologists term denudation rates: these are rates of mass loss, or of equivalent surface lowering, from landscapes as a result of physical erosion and chemical weathering. Long-term denudation rates can be inferred from cosmogenic nuclides — rare isotopes that are produced inside mineral grains by cosmic-ray bombardment<sup>3</sup>. Measuring the concentrations of these nuclides is technically demanding, but the effort is worthwhile because they intrinsically reflect average denudation rates over thousands or tens of thousands of years. And when they are measured in handfuls of river sand, they reflect the average denudation rate of the entire contributing drainage basin. In the nearly 20 years since this technique was proven to work<sup>4</sup>, more than 1,000 such measurements of basin-scale denudation rates have been published.

Using their global compilation of cosmogenic-nuclide measurements, Willenbring *et al.* show that roughly half of the worldwide variation in denudation rates can be explained by a simple exponential function of average basin slope. This implies that other factors such as climate, rock type and vegetation must be of secondary importance. Another recent compilation of cosmogenic-nuclide data<sup>5</sup> is consistent with this exponential relationship, as are other, smaller data sets covering different spatial scales, timescales and measurement methods<sup>4,6,7</sup>. This exponential relationship implies that denudation rates do not decline to zero as slope approaches zero; instead, they converge to a constant value of roughly 0.01–0.1 millimetres per year.

These rates are slower than those in steep mountainous terrain by an order of magnitude or more, but gentle landscapes are so abundant that they dominate the global denudation budget. Willenbring and co-workers calculate that landscapes with slopes of less than 100 metres per kilometre account for more than nine-tenths of Earth's land surface and eight-tenths of the global denudation flux, even when the faster rates in steeper terrain are taken into account. The implication is that understanding how Earth's topography and climate have co-evolved over geological time — and how they will co-evolve in the future — will require much better data on erosion and weathering processes in flat terrain.

Getting those data will be a challenge. We



**Figure 1 | Lowland flat, mountain high.** Flat terrain surrounds towering peaks in the eastern Tibetan Plateau. Willenbring and colleagues' analysis<sup>1</sup> suggests that geologists need to focus on gently sloping terrain, rather than mountains, if we are to understand the denudation processes that transform Earth's surface.

currently have far fewer measurements of long-term denudation rates in gentle terrain than in steep terrain<sup>8</sup>. Not one of the nearly 1,000 denudation-rate measurements in Willenbring and colleagues' database comes from a landscape with a slope of less than 10 m km<sup>-1</sup>, even though such landscapes account for roughly half of Earth's land surface<sup>1</sup>.

There are details in the paper that one can argue over. For example, measurements of topographic slope are scale-dependent; widely spaced elevation measurements straddle valleys and ridges and thus underestimate topographic gradients, potentially by a lot. Therefore, Earth may not be nearly as flat as one would infer from the global topographic data set used by the authors, which has a spatial resolution of only 1 km.

More fundamentally, a lot of gentle terrain consists of depositional landforms such as basins, deltas, fans and playas, where sediment accumulates, rather than erodes, over the long term. Cosmogenic nuclides can say nothing useful about the denudation rates of depositional landscapes (which are, strictly speaking, negative). Therefore, cosmogenic nuclides are typically measured only in the subset of gentle landscapes that are actively eroding, and not in those that accumulate sediment. This will lead to a systematic upward bias in the inferred average denudation rate of gentle terrain, and so to an upward bias in the importance of gentle terrain in Willenbring and colleagues' global denudation budget. The magnitude of this bias is unknown.

Nonetheless, the authors' analysis does suggest that more geologists should climb down from the mountains that they love so much, to explore the erosion and weathering of the surrounding lowlands (Fig. 1). When they get there, however, they will find that many of these landscapes have been utterly

transformed by human activities, which have greatly accelerated soil-erosion rates — to the point that they rival or exceed natural erosion rates in some of the steepest mountains on Earth, and threaten the long-term future of our food supply<sup>8</sup>.

This is not a new problem: the crumbling relics of many vanished civilizations bear mute witness to the fate of those who squander their soil resources<sup>9</sup>. What is new, this time around, is that we know we are destroying our soil, and we know how to stop. The scientific basis for soil conservation has been understood for decades. What remains to be discovered, in many cultural landscapes, is the will to implement it. ■

**James W. Kirchner** is in the Department of Environmental System Sciences, ETH Zurich, CH-8092 Zurich, Switzerland, and at the Swiss Federal Research Institute WSL, Birmensdorf, Switzerland. **Ken L. Ferrier** is in the Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138, USA. e-mails: kirchner@ethz.ch; ferrier@fas.harvard.edu

- Willenbring, J. K., Codilean, A. T. & McElroy, B. *Geology* **41**, 343–346 (2013).
- Riebe, C. S., Kirchner, J. W. & Finkel, R. C. *Earth Planet. Sci. Lett.* **224**, 547–562 (2004).
- Granger, D. E. & Riebe, C. S. in *Treatise on Geochemistry* Vol. 5 (eds Holland, H. D. & Turekian, K. K.) 1–43 (Elsevier, 2007).
- Granger, D. E., Kirchner, J. W. & Finkel, R. C. *J. Geol.* **104**, 249–257 (1996).
- Portenga, E. W. & Bierman, P. R. *GSA Today* **21**, 4–10 (2011).
- Montgomery, D. R. & Brandon, M. T. *Earth Planet. Sci. Lett.* **201**, 481–489 (2002).
- Summerfield, M. A. & Hulton, N. J. *J. Geophys. Res. (Solid Earth)* **99**, 13871–13883 (1994).
- Montgomery, D. R. *Proc. Natl Acad. Sci. USA* **104**, 13268–13272 (2007).
- Montgomery, D. R. *Dirt: The Erosion of Civilizations* (Univ. California, 2012).

## CANCER

# Trouble upstream

**Mutations in the promoter sequence of the gene encoding a component of the enzyme telomerase have been discovered in patients with melanoma, suggesting that mutations in this regulatory region can drive tumour development.**

E. ELIZABETH PATTON & LEA HARRINGTON

One of the hallmarks of cancer is the increased production of telomerase, an enzyme that replenishes chromosome ends and enables continued cell division. But how and when does production of telomerase become selected for in the evolution of a tumour? Writing in *Science*, Huang *et al.*<sup>1</sup> and Horn *et al.*<sup>2</sup> show that subtle genetic changes in the promoter region that lies upstream of the gene encoding the telomerase subunit TERT (telomerase reverse transcriptase) are driver mutations in melanoma, and possibly in other cancers.

The two groups arrived at this discovery from different perspectives. Huang and colleagues analysed whole-genome sequence data for cells taken from human malignant melanomas and focused on the DNA that was not part of protein-coding sequences. In more than 70% of these melanoma samples, they found one of two single-nucleotide C-to-T mutations

or a dinucleotide CC-to-TT substitution in the core promoter region of *TERT* (Fig. 1). All three mutations are suggestive of damage induced by ultraviolet (UV) light; UV light is the primary environmental risk factor for melanoma, and an abundance of such 'UV signature' mutations is known in this skin cancer<sup>3-5</sup>. The authors also identified these *TERT*-promoter mutations at a frequency of about 16% in a panel of cell lines derived from a range of other cancers, with a somewhat higher than average incidence in bladder and liver cancers.

Horn and colleagues began their study by examining the genetics of a large family to find a germline (inherited) mutation that predisposed four generations of family members to melanoma. Using gene linkage analysis followed by targeted sequencing, they identified multiple sequence variants linked to the disease, one of which was a T-to-G mutation in the *TERT* promoter. The researchers then screened melanoma cells from unrelated patients for other examples of this mutation,

but they instead found a high frequency of UV-signature mutations in the *TERT* promoter, including those identified by Huang *et al.* and an additional CC-to-TT double substitution.

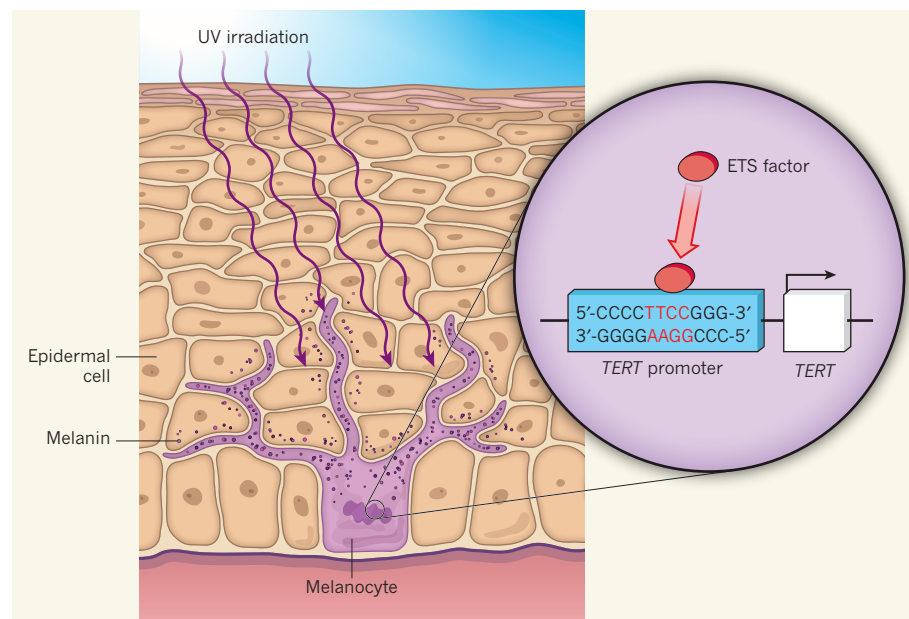
Until now, *BRAF* has been the gene most frequently identified as mutated in melanomas and nevi (commonly known as moles), but the combined frequency of *TERT* mutations reported in the present studies is even higher than for *BRAF*. Interestingly, Horn *et al.* did not find the *TERT*-promoter mutations to be a common feature in naevi, although there is insufficient evidence to indicate whether they are associated with the progression of benign naevi to melanoma.

Frequent *TERT*-promoter mutations in somatic (non-germline) cells have not been identified as driver mutations in cancers before, and the genome of melanoma cells has a notoriously high somatic mutation rate. So it was important for both groups to show a functional relevance of the mutations and to provide evidence that the mutations did not arise by chance.

A key finding supporting such relevance is that all the mutations identified have one feature in common: they generate a sequence that can be bound by ETS transcription factors (Fig. 1), and thus have the potential to confer additional transcriptional control on the *TERT* promoter. Indeed, this seems to be the case — the authors observed a two- to four-fold increase in *TERT* expression when they tested the mutated promoter sequence in a gene-expression 'reporter system', compared with the wild-type promoter. Further experiments will be required to determine whether this increase is also observed in patients. Additionally, neither group found the mutations together in the same melanoma, implying that the mutations are functionally redundant and may each confer a selective advantage, rather than occurring by chance owing to a locally elevated mutation rate.

Can a subtle change in *TERT*-gene expression give certain cancer cells a growth advantage? Telomeres are non-coding sequences at the ends of chromosomes that shorten with every cell division as a result of incomplete DNA replication; telomerase acts to replenish these sequences. When the amount of functional telomerase is reduced, even by only half, telomere erosion occurs in yeast, mice and humans<sup>6</sup>. In humans and mice, mutations in *TERT* that result in only a single functional copy of the gene adversely affect the function of tissues that have a high cell turnover, such as skin and blood, and can cause early mortality<sup>5</sup>. Conversely, increased telomerase production can promote cancer progression; a dramatic example of this is the malignant progression of prostate cancer that occurs when telomerase activity is restored in a telomerase-negative, tumour-prone mouse model<sup>7</sup>.

Notably, it seems that melanocytes (the



**Figure 1 | Promoter mutations create new transcription-factor binding sites.** Cells called melanocytes reside in the epidermal layer of the skin and provide surrounding skin cells with the pigment melanin to protect them from the damaging effects of ultraviolet (UV) light. Huang *et al.*<sup>1</sup> and Horn *et al.*<sup>2</sup> identify 'UV signature' mutations in the promoter region of the gene encoding TERT, a subunit of the enzyme telomerase. Strikingly, four mutations identified by the authors all result in nucleotide sequences that contain a binding site (red) for transcription factors of the ETS family. Binding of these proteins leads to increased TERT expression. Horn *et al.* also describe a UV-independent germline mutation in the *TERT* promoter in cells from members of a melanoma-prone family that similarly generates a binding site for ETS transcription factors.



skin cells in which melanoma arises) are especially sensitive to changes in *TERT* expression, because in Horn and colleagues' study all but one person carrying the T-to-G mutation developed melanoma, and the unaffected carrier had multiple naevi. This susceptibility could arise because ETS proteins are targets of the MAPK cell-signalling pathway, which is frequently activated in melanoma and benign naevi. In support of this idea, Horn and colleagues found that the *TERT*-promoter mutations were found in many of the melanomas that had activating *BRAF* mutations; *BRAF* is an intermediate in the MAPK pathway, and perhaps increased *BRAF* signalling leads to enhanced activation of ETS factors at the *TERT* promoter.

Another, non-mutually exclusive possibility is that tissue-specific expression of ETS proteins causes some cell types to become particularly vulnerable to the effects of *TERT*-promoter mutations. Indeed, ETS transcription factors are expressed in melanomas<sup>8</sup>. Furthermore, the T-to-G germline mutation identified by Horn *et al.* generates a binding site that may be more selective for Elk1 and Elk4, a subgroup of ETS that is highly expressed in the ovary, and the authors report that two women with this *TERT*-promoter mutation also developed ovarian cancer.

These studies reveal two conceptual advances. First, they represent an unusual type of driver mutation in cancer — one that need not modify the gene product itself, but instead creates a binding site for transcriptional activators that leads to an increase in normal gene expression. Second, they show that changes in *TERT* expression may be sufficient to cooperate with cancer-initiating oncogenes to drive cancers such as melanoma. A crucial next experiment will be to examine the functional consequences of the increase in *TERT* expression in patients with *TERT*-promoter mutations, to see whether it confers a survival advantage through a decrease in the incidence of short or fused telomeres that often persist in dividing cancer cells<sup>9,10</sup>. Taken together, the findings suggest that cancer development can be driven by relatively subtle changes in the expression level of a normal gene, indicating that trouble can lie upstream of genes as well as within them. ■

**E. Elizabeth Patton** is at the MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XR, UK.

**Lea Harrington** is at the Wellcome Trust Centre for Cell Biology, Edinburgh EH9 3JR, UK, and the Institute for Research in Immunology and Cancer, University of Montreal, Canada. e-mails: e.patton@igmme.ed.ac.uk; l.harrington@ed.ac.uk

- Huang, F. W. *et al. Science* **339**, 957–959 (2013).
- Horn, S. *et al. Science* **339**, 959–961 (2013).
- Pleasance, E. D. *et al. Nature* **463**, 191–196 (2010).

- Hodis, E. *et al. Cell* **150**, 251–263 (2012).
- Krauthammer, M. *et al. Nature Genet.* **44**, 1006–1014 (2012).
- Harrington, L. *Mutat. Res.* **730**, 37–42 (2012).
- Ding, Z. *et al. Cell* **148**, 896–907 (2012).

- Jané-Valbuena, J. *et al. Cancer Res.* **70**, 2075–2084 (2010).
- Capper, R. *et al. Genes Dev.* **21**, 2495–2508 (2007).
- Xu, L. & Blackburn, E. H. *Mol. Cell* **28**, 315–327 (2007).

## CONDENSED-MATTER PHYSICS

# A frustrated trio

**Geometrical frustration results from an incompatibility between the spatial arrangement of a system's constituents and the interactions between them. The effect has now been observed in a triangular triple quantum dot.**

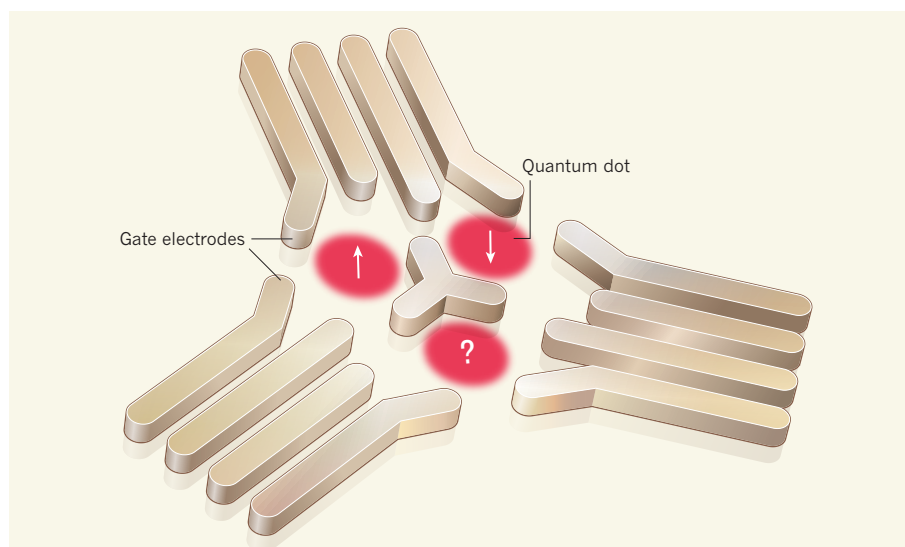
SABINE ANDERGASSEN

**F**rustration arises when competing interactions cannot be satisfied simultaneously. This compromise is central to the behaviour of many complex systems, from neural networks to folding proteins and magnetic materials. Writing in *Physical Review Letters*, Seo *et al.*<sup>1</sup> describe experiments on a simple frustrated system that make a promising step towards our understanding of the mechanisms governing the physics of frustration.

In condensed-matter physics, frustration is a feature of magnetic materials that have competing interactions between localized magnetic moments (spins). In a ferromagnet, the spins align with one another and result in a macroscopic magnetic moment. By contrast, in an antiferromagnet, neighbouring spins point in opposite directions and the system exhibits no such macroscopic moment. An

interesting situation arises when the system's lattice geometry inhibits the formation of an ordered, low-temperature spin configuration. The origin of this effect is remarkably simple, and can be illustrated by three spins interacting antiferromagnetically on a triangular lattice. Once two of the spins on an elementary triangle are anti-aligned, the third one is 'frustrated': it can no longer point in a direction opposite to both of the other spins. This incompatibility of antiferromagnetic interactions with the underlying lattice geometry is known as geometrical frustration, and is particularly pronounced in low dimensions, in which the effects of interactions prevail.

The simple, and at the same time, fundamental model of geometrical frustration in a triangle exhibits six competing lowest-energy spin configurations with the same energy (degenerate ground states). The resulting fluctuations between these states suppress



**Figure 1 | Geometrical frustration in a triangle of three quantum dots.** In a triangular triple quantum dot in which the dots' electron spins (arrows) interact antiferromagnetically, if two of the quantum dots are in the 'up' and 'down' spin states, respectively, owing to the triangular geometry of the system, the spin of the third quantum dot is frustrated as it can align opposite to only one of the other spins. As a result, there are six configurations (all possible combinations of spin-up and spin-down except for the two in which all three spins align) with the same energy. By controlling the system with gate electrodes, Seo and colleagues<sup>1</sup> observe this sixfold degeneracy.

conventional ordering and the system features non-zero entropy even at zero temperature, in contrast to the naive expectation from the third law of thermodynamics. Quantum superposition of degenerate states leads to the emergence of exotic many-body phenomena<sup>2–6</sup>, from spin ice and quantum spin liquids to high-temperature superconductivity.

A key characteristic of frustrated systems is a large number of degenerate ground states. This high degeneracy is very sensitive to perturbations: any slight asymmetry in the system will reduce the number of degenerate ground states. This instability leads to rich physical behaviour. Understanding the fundamental mechanisms and principles underlying the variety of quantum phenomena that arise from highly degenerate ground states is of conceptual importance in condensed-matter physics. Both the development of numerical-simulation techniques and the increasing ability to experimentally engineer systems in which magnetic, charge or vibrational degrees of freedom interact, have a central role in addressing this problem. In particular, optical lattices (arrays of interfering laser beams) of cold atoms<sup>7–9</sup> provide model systems for strongly interacting many-body systems.

Quantum dots (artificial atoms) provide a promising platform for investigating non-equilibrium charge transport, and in particular how it is affected by geometrical frustration. Advanced techniques for fabricating nanoscale devices allow detailed modelling and flexible tuning of parameters. In semiconductor quantum dots, the potential in which the dots are confined and the number of electrons, as well as the interaction between them, are controlled by metallic 'gate' electrodes.

In their study, Seo *et al.* report the first observation of geometrical frustration in quantum dots. Because of the difficulties in building highly symmetrical devices, this had not been achievable previously. Their frustrated quantum system consists of three quantum dots arranged in a triangle (Fig. 1). Antiferromagnetic interactions between the dots' electron spins lead to a frustrated ground state, which results from the competition of the six degenerate, three-spin configurations out of eight possible spin arrangements. Because controlling electron spins in quantum dots is difficult, the authors used an alternative way of studying geometrical frustration: degenerate charge states. The advantage of this approach is that these charge states (isospins) are defined by the parity of the number of charge carriers in the dots and can be precisely controlled by metallic gate electrodes. The authors' measurements of charge transport revealed the six-fold degeneracy of the isospin configurations in the maximum electrical conductance induced by the fluctuations between the states.

Seo and colleagues' work sheds light on the impact of frustration on charge transport and its characteristic signatures, and motivates

further experimental as well as theoretical studies, including analysis of the time evolution of frustrated quantum systems. Quantum dots have been suggested as potential platforms for implementing spin or charge quantum bits<sup>10–14</sup> for quantum computation and information processing. Control of magnetism on the atomic scale is also becoming essential as data-storage devices are miniaturized. In particular, switchable nanoscale antiferromagnets are being discussed as candidate building blocks for future memory, storage and 'spintronic' applications<sup>15</sup>. Understanding the underlying physics will be crucial for the development of these technologies. The challenge is the scalability to large systems — to explore complex many-body phenomena, identify novel quantum phases and design interesting quantum materials. ■

**Sabine Andergassen** is in the Faculty of Physics, University of Vienna, 1090 Vienna, Austria.

#### MOLECULAR BIOLOGY

## Circles reshape the RNA world

**The versatility of RNA seems limitless. The latest surprise comes from circular RNAs, which are found to counteract the function of another class of regulatory RNA — the microRNAs. SEE ARTICLE P.333 & LETTER P.384**

**KENNETH S. KOSIK**

**T**he protein-coding function of messenger RNAs can be suppressed by the binding of short microRNA sequences. But how microRNA-induced suppression is itself inhibited is poorly understood. In this issue, Memczak *et al.*<sup>1</sup> (page 333) and Hansen *et al.*<sup>2</sup> (page 384) describe highly stable, circular RNAs that bind several copies of a microRNA to terminate suppression of mRNA targets\*.

The circular RNA (circRNA) reported, called CDR1as by Memczak *et al.* and ciRS-7 by Hansen *et al.*, contains roughly 70 evolutionarily conserved binding sites for microRNA-7 (miR-7) and forms a complex with AGO proteins. The latter are part of the RNA-induced silencing complex, which allows miRNAs to recognize their target mRNAs. When Memczak and colleagues expressed human CDR1as/ciRS-7 in zebrafish embryos, its effects were the same as those seen when miR-7 expression was reduced — impaired midbrain development. Moreover, the authors' bioinformatic predictions indicated that

\*This article and the papers under discussion<sup>1,2</sup> were published online on 27 February 2013.

e-mail: [sabine.andergassen@univie.ac.at](mailto:sabine.andergassen@univie.ac.at)

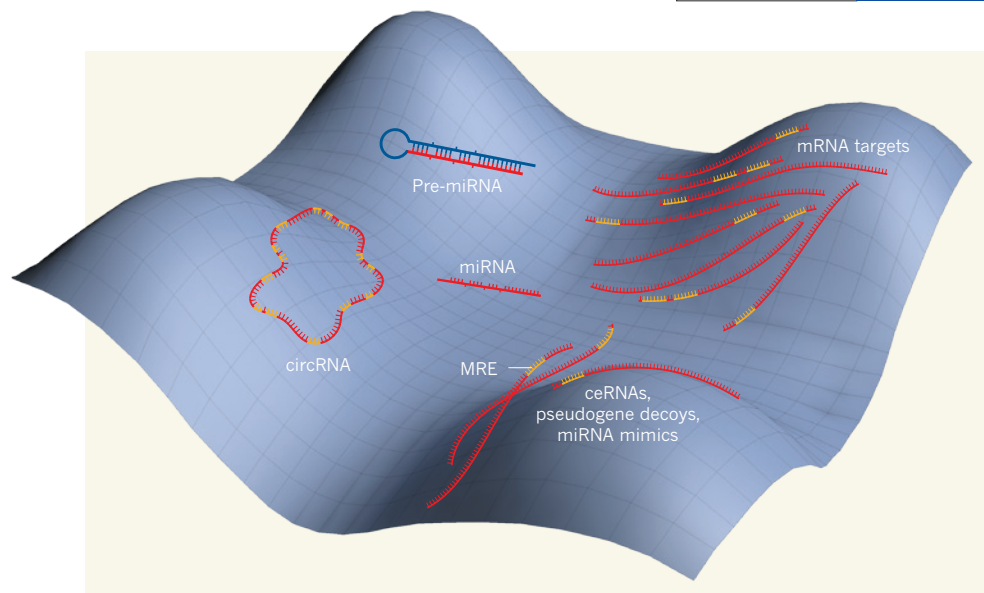
1. Seo, M. *et al.* *Phys. Rev. Lett.* **110**, 046803 (2013).
2. Bramwell, S. T. & Gingras, M. J. P. *Science* **294**, 1495–1501 (2001).
3. Ramirez, A. *Nature* **399**, 527–528 (1999).
4. Balents, L. *Nature* **464**, 199–208 (2010).
5. Pratt, F. *et al.* *Nature* **471**, 612–616 (2011).
6. Bramwell, S. T. *et al.* *Nature* **461**, 956–959 (2009).
7. Bloch, I. *Nature* **453**, 1016–1022 (2008).
8. Bloch, I., Dalibard, J. & Zwerger, W. *Rev. Mod. Phys.* **80**, 885–964 (2008).
9. Bloch, I., Dalibard, J. & Nascimbène, S. *Nature Phys.* **8**, 267–276 (2012).
10. Loss, D. & DiVincenzo, D. P. *Phys. Rev. A* **57**, 120–126 (1998).
11. Petta, J. R. *et al.* *Science* **309**, 2180–2184 (2005).
12. Koppens, F. H. L. *et al.* *Nature* **442**, 766–771 (2006).
13. Hanson, R., Kouwenhoven, L. P., Petta, J. R., Tarucha, S. & Vandersypen, L. M. K. *Rev. Mod. Phys.* **79**, 1217–1265 (2007).
14. Pioro-Ladrière, M. *et al.* *Nature Phys.* **4**, 776–779 (2008).
15. Loth, S., Baumann, S., Lutz, C. P., Eigler, D. M. & Heinrich, A. J. *Science* **335**, 196–199 (2012).

thousands of circRNAs reside in the genome, consistent with previous reports<sup>3,4</sup>.

Target suppression by miRNAs is highly nuanced. On the one hand, these sequences can induce AGO-mediated endonucleolytic mRNA cleavage triggered by complementarity between the mRNA and the miRNA at nucleotides 10 and 11. The destruction of the target, which follows, frees the miRNA to bind to its next target in a catalytic manner. On the other hand, miRNAs can inhibit protein translation by binding more stably to a target mRNA in a stoichiometric manner. This makes the target a 'reservoir' that prevents the miRNA from inhibiting other mRNA targets. The latter mechanism is an indication of the way in which competing endogenous RNAs (ceRNAs) act. These are mRNAs that share miRNA-response elements (MREs) with other mRNAs and so compete for binding to those miRNAs with which they also share MREs<sup>5</sup>.

Like ceRNAs, circRNAs serve as miRNA reservoirs. However, circRNAs have numerous binding sites for a specific miRNA and so are completely dedicated to their role of harbouring miRNAs. Binding of a miRNA to a ceRNA not only prevents that miRNA from binding to





**Figure 1 | Constraints on evolutionary change in microRNAs.** MicroRNAs (miRNAs) lie in a fitness valley constrained by their numerous interactions, which include those with the hairpin structure of the precursor miRNA (pre-miRNA), the many target mRNAs and other RNAs that terminate or modulate miRNA binding to target sequences by competing against them. The latter category includes competing endogenous RNAs (ceRNAs), pseudogene decoys and miRNA mimics. Two studies<sup>1,2</sup> introduce circular RNAs (circRNAs) as another constraining factor. MRE, miRNA-response element.

other MREs, but can also suppress translation from the coding portion of the ceRNA. Hence, compared with circRNAs, ceRNAs operate in a more complex weave of interacting molecules that constrains translation. Other reservoirs of target sites also reside on distinct molecules. These include target mimics such as the *IPS1* gene in the plant *Arabidopsis thaliana*<sup>6</sup>, decoys within pseudogenes such as *PTENP1* (ref. 7) and possibly 3'-untranslated regions of mRNA that are expressed separately from their associated protein-coding sequences<sup>8</sup>.

Circularizing RNA enhances its stability by obviating a role for RNA exonuclease enzymes, which act on free 3' and 5' ends of an RNA molecule to cleave it. Moreover, with several binding sites dedicated to antagonizing a single miRNA, a circRNA can capture miRNAs from numerous targets in one fell swoop. Likewise, circRNA destruction could release a shower of miRNAs that target multiple mRNAs with the shared MRE. In fact, Hansen *et al.* outline a circRNA-destruction mechanism in which miR-671 binds CDR1as/ciRS-7 with greater complementarity than miR-7 and induces AGO-mediated cleavage of this circRNA.

Snapshot approaches to profiling miRNAs reveal that the greatest changes in their expression levels occur at transition points in development, cell differentiation or carcinogenesis<sup>9</sup>. Clearance of the mRNA-miRNA duplexes at these points and their replacement with different miRNAs could operate through circRNAs. For instance, as a brute means of vacuuming up miRNAs, circRNAs could increase in expression as cell differentiation from stem cells proceeds, to capture the exceedingly high levels of miRNAs expressed in stem cells. They

could also clean up the opposite strands of mature miRNAs, which can be present in surprisingly large numbers<sup>10</sup>, or potentially function therapeutically to divert cancer-associated miRNAs from promoting an oncogenic pathway. In all these cases, however, the circRNAs sequester miRNAs, and so a knowledge gap remains regarding how miRNAs are destroyed.

To function optimally, the number of miRNA-binding sites on each circRNA is probably under selection pressure to attract nearly all of a specific miRNA population from all of its target sites. If so, the number of miRNA-binding sites on a circRNA multiplied by the number of copies of the circRNA in a single cell will inform us about the collective strength of all MREs for a particular miRNA. Many miRNAs operate at copy numbers of  $10^3$  per cell — a likely lower boundary for the number of circRNA sites required to mop them all up. However, for circRNAs to win out against mRNA targets in the competition for miRNA binding, they must have a greater affinity for the miRNAs. High affinity can be thermodynamically built into the circRNA sequence, but may also require an excess of circRNA-encoded miRNA-binding sites relative to the total number of other relevant MREs in the cell. Certainly, modellers will soon be romping through this territory.

The fitness 'landscape' that has contributed to maintaining each of the roughly 21-nucleotide miRNAs unchanged over major parts of evolution includes circRNAs (Fig. 1). The selection pressure on each miRNA nucleotide is undoubtedly high: an miRNA sequence must base-pair to itself to form the hairpin-shaped precursor miRNA; it must pair with a host of target mRNAs; and it must pair with

binding sites that terminate or modulate target interaction. Despite the enormous number of possible miRNA sequences, the small amount of change in miRNAs implies that the remaining evolutionary space for innovation is limited; in other words, miRNAs have approached molecular perfection. Throughout animal evolution, nature has tinkered with the sequences of a relatively constant set of coding genes, whereas miRNA innovation, in general, is more reliant on the invention of completely novel sequences<sup>11</sup>. Perhaps the ease with which hairpin-shaped miRNA precursors can arise as potential regulatory elements — and fit ‘digitally’ into a wealth of genomic non-coding sequence, including circRNAs — could serve as a driver of evolution.

As a footnote, a better naming system for circRNAs is needed. ‘ciRS-7’ denotes binding to miR-7, and therefore assumes that other circRNAs in this category will also neatly align with a single miRNA. ‘CDRIas’ assumes that circRNAs will bear some relationship to a named gene — in this case, an antisense sequence to the cerebellar degeneration-related gene. With thousands of these circRNAs in the genome, they require their own numbering system. My suggestion is that this one is called circR-1. ■

**Kenneth S. Kosik** is at the Neuroscience Research Institute and in the Department of Molecular, Cellular and Developmental Biology, University of California,

Santa Barbara, California 93106, USA.  
e-mail: kosik@lifesci.ucsb.edu

1. Memczak, S. *et al.* **495**, 333–338 (2013).
2. Hansen, T. B. *et al.* *Nature* **495**, 384–388 (2013).
3. Danan, M., Schwartz, S., Edelheit, S. & Sorek, R. *Nucleic Acids Res.* **40**, 3131–3142 (2012).
4. Salzman, J., Gawad, C., Wang, P. L., Lacayo, N. & Brown, P. O. *PLoS ONE* **7**, e30733 (2012).
5. Tay, Y. *et al.* *Cell* **147**, 344–357 (2011).
6. Franco-Zorrilla, J. M. *et al.* *Nature Genet.* **39**, 1033–1038 (2007).
7. Poliseno, L. *et al.* *Nature* **465**, 1033–1038 (2010).
8. Mercer, T. R. *et al.* *Nucleic Acids Res.* **39**, 2393–2403 (2011).
9. Neveu, P. *et al.* *Cell Stem Cell* **7**, 671–681 (2010).
10. Zhou, H. *et al.* *Nucleic Acids Res.* **40**, 5864–5875 (2012).
11. Tarver, J. E., Donoghue, P. C. & Peterson, K. J. *BioEssays* **34**, 857–866 (2012).

to the general process of galaxy formation and evolution. Ambitious wide-field surveys from the ground, ranging from the Sloan Digital Sky Survey to the recently started Dark Energy Survey, are also driving our view of the properties of galaxies, and enabling unprecedentedly accurate and complete statistical studies.

A fresh opportunity to discover unusual galaxies, magnified by the gravitational influence of foreground objects, has just been provided by a combination of the wide-area, millimetre-wavelength surveys obtained with the South Pole Telescope (SPT) and the imaging and spectroscopic power of ALMA at the galaxies’ pinpointed locations. Vieira and colleagues take this opportunity in their study. Their results come from an SPT survey of 1,300 square degrees, an area of sky that blocks out about the same angle as that subtended by a laptop screen in front of its user. The SPT takes images of the sky at three wavebands in the millimetre range, with the primary objective of identifying distant clusters of galaxies by means of the Sunyaev-Zeldovich effect — the characteristic spectral signal imprinted in these bands on the cosmic microwave background relic radiation from the Big Bang, due to the scattering of this radiation by hot electrons in the clusters<sup>4</sup>.

## ASTRONOMY

# The ALMA telescope shows its true colours

**Bright, gravitationally magnified galaxies have been found across a wide span of cosmic time. The first results from the still-growing ALMA telescope show its power to reveal these galaxies’ redshifts and internal structure. SEE LETTER P.344**

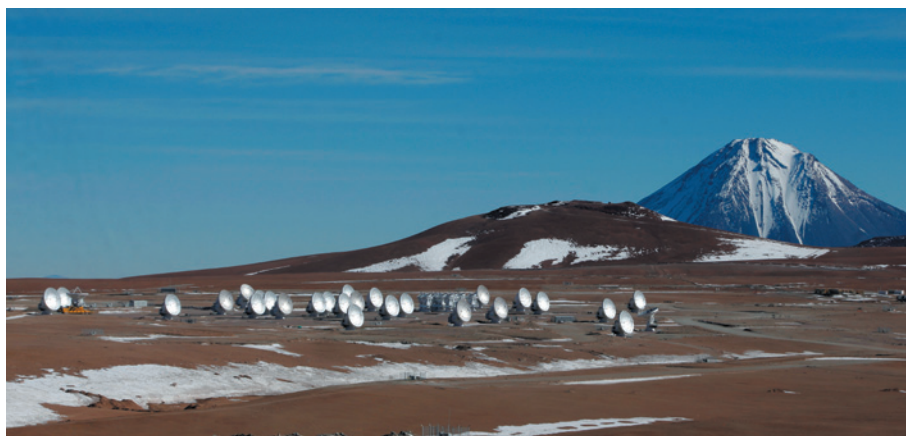
ANDREW W. BLAIN

The Atacama Large Millimeter/submillimetre Array (ALMA) interferometer, which is nearing completion in Chile<sup>1</sup>, is revolutionizing observational astronomy. It provides precise views of the internal workings of galaxies by imaging the gas and dust in their interstellar medium at wavelengths of 0.3–3 millimetres. On page 344 of this issue, Vieira *et al.*<sup>2</sup> present some of the first detailed images and measurements of redshifts of distant galaxies obtained by ALMA (Fig. 1). The authors targeted galaxies discovered in a wide-field survey using the South Pole Telescope<sup>3</sup>, to find those that are magnified — and so made unusually easy to study — by the ‘gravitational lensing’ effect of foreground objects\*.

The steady discovery of new classes of galaxy, as different windows on the cosmos have become accessible, has been a highlight of astrophysics for decades. The radio, X-ray and far-infrared windows have opened to reveal ever more about the processes at work in both long-known and previously unappreciated types of galaxy. By combining the power of wide-field surveys to discover galaxies and the ability to dissect the galaxies’ nature using high-resolution imaging and spectroscopy, an improved understanding of their properties can be obtained rapidly and efficiently.

\*This article and the paper under discussion<sup>2</sup> were published online on 13 March 2013.

The heritage of this process dates back to the 1920s with the spectroscopy of diffuse nebulae obtained using the largest optical telescope then available, which led to the measurement of the expansion rate of the Universe. Then came the recognition of radio sources and other active galactic nuclei (AGN) in the 1960s, unambiguously showing cosmic evolution taking place. Wide-field, space-based surveys using the IRAS infrared and ROSAT X-ray satellites several decades ago, and more recently the far-infrared Herschel Space Observatory, mid-infrared WISE and millimetre-wavelength Planck spacecraft, have found classes of unusual and rare galaxies, and clear clues



**Figure 1 | The ALMA array in the Atacama Desert of northern Chile.**

C. PADILLA/ALMA (ESO/NAOJ/NRAO)



In parallel, the SPT also detects a bright sample of far-infrared-luminous galaxies, which are shining because their interstellar medium absorbs visible and ultraviolet light from stars and AGN. Their measured colours in the SPT images are used to weed out galaxies with strong radio emission from AGN, and galaxies that are sufficiently near to us to have been catalogued using the IRAS satellite. This leaves an efficiently selected, reliable catalogue of distant, bright galaxies, which are typically magnified by the gravitational lensing (or light-bending) effect of foreground objects<sup>5</sup>. The astrophysics of these galaxies can be studied in great detail, if a suitable tool is available to detect and resolve their internal structure<sup>6</sup>, by taking advantage of the lensing to increase both their apparent size on the sky and their total emitted power. Careful modelling can provide a view of details that would be beyond the reach of even weeks of observations without the effect of this 'gravitational telescope'.

Vieira *et al.* showcase the power of ALMA by making images and spectra of 47 galaxies detected in their SPT survey. They use two techniques: imaging at relatively short, 1-mm wavelengths to reveal the 'sheared' and multiply imaged views of the galaxies formed by gravitational lensing, and spectral scanning across the ALMA waveband at around 3 mm, to find the galaxies' redshifts from line emission coming from their interstellar-medium gas. In a large fraction (about 88%) of the 26 spectral-scanning observations, secure serendipitous redshifts were found from multiple spectral lines. The measured redshift range (2–5.7) corresponds to a substantial fraction of the history of the Universe, with the earliest objects dating back to one billion years after the Big Bang. Furthermore, the forthcoming ability of ALMA to measure the internal motion of the gas in a substantial sample of distant galaxies, using spatially resolved spectral measurements, to reveal their shapes, sizes, rotation speeds, masses and any merging subunits, is unprecedented. By exploiting lensing, the time required to detect and resolve these targets using ALMA is reduced to minutes, which is ideally suited to exploit the precious time available for science as the instrument is brought into full service. The lensing effect may subtly bias the sample towards higher redshifts, but this can be addressed by careful analysis of larger samples from the ongoing SPT surveys<sup>3</sup>.

There is now a steady stream of distant, far-infrared-luminous, dust-enshrouded galaxies that ALMA can target. The fruits of galaxy surveys by the Herschel Space Observatory, now entering the last months of its multi-year mission, are available<sup>7</sup>, as are those from all-sky surveys made using the Planck and WISE missions, to catch the most extreme examples of the galaxy population in the act of formation. New facilities on the ground, including the SCUBA-2 camera on the James Clerk

Maxwell Telescope in Hawaii<sup>8</sup> and new cameras becoming available at the IRAM 30-m telescope in Spain, are surveying at shorter wavelengths than the SPT, bridging the gap between the millimetre-wavelength SPT and the far-infrared Herschel. These observations will yield additional targets, and the Atacama Cosmology Telescope in Chile<sup>9</sup> is also conducting a suite of complementary millimetre-wavelength surveys. The sensitivity and wide field of view of the planned CCAT telescope<sup>10</sup> will ultimately advance the precision and resolution of wide-field far-infrared/submillimetre-band surveys well beyond these current survey facilities.

The promise of ALMA for unravelling the findings of wide-field dusty-galaxy surveys, as conducted by the SPT, has now been clearly demonstrated. It is exciting to see the results, but they are a mere prelude to the final capabilities of this awesome machine — the first global-scale international observatory that has been constructed by a genuinely equal partnership between parties on different

continents. With a spatial resolution finer than that of the Hubble Space Telescope and an 'always-on' spectroscopic capability, ALMA can probe an otherwise invisible fraction of the energy emitted by galaxies, to provide a much more complete view of galaxy dynamics and evolution, and the processes that form stars and fuel AGN. ■

**Andrew W. Blain** is in the University of Leicester, Department of Physics and Astronomy, Leicester LE1 7RH, UK.  
e-mail: ab520@star.le.ac.uk

1. www.almaobservatory.org
2. Vieira, J. D. *et al.* *Nature* **495**, 344–347 (2013).
3. <http://pole.uchicago.edu>
4. Reichardt, C. L. *et al.* Preprint at <http://arxiv.org/abs/1203.5775> (2012).
5. Vieira, J. D. *et al.* *Astrophys. J.* **719**, 763–783 (2010).
6. Swinbank, A. M. *et al.* *Astrophys. J.* **742**, 11 (2011).
7. Negrello, M. *et al.* *Science* **330**, 800–804 (2010).
8. Chen, C.-C. *et al.* *Astrophys. J.* **762**, 81 (2013).
9. www.princeton.edu/act
10. www.ccatobservatory.org

#### EVOLUTIONARY GENOMICS

## Detecting selection

**Advances in population genetics and genome sequencing have made it possible to identify anonymous fragments of DNA that have undergone selection. This yields some evolutionary answers, and a panoply of puzzles. SEE LETTER P.360**

GREGORY S. BARSH & LEIF ANDERSSON

**W**hen an affenpinscher named Banana Joe took the top honours at last month's Westminster Kennel Club dog show in New York, the judge referred to his "fantastic face [and] great body". Some of Joey's success may be attributed to the awesome power of artificial selection, but we sometimes overlook the fact that dogs are also our companions in natural selection, and have adapted to similar changes in lifestyle and nutrient availability over the past 10,000 years or so.

The parallel evolution of humans and our animal companions is apparent in a study by Axelsson *et al.*<sup>1</sup> on page 360 of this issue, in which the authors used a population-genomics approach to identify regions of the dog genome that have undergone selection during domestication. This method promises to revolutionize evolutionary biology, by challenging us to detect traits affected by evolution on the basis of genotype rather than an organism's characteristics, or phenotype. Two other recent papers, published in *Cell* by Grossman *et al.*<sup>2</sup> and Kamberov *et al.*<sup>3</sup>, rise to this challenge and show how hypotheses about an adaptive human genotype can be tested in controlled

experiments. Together, the three papers are a wonderful intersection between genomics, population science and experimental genetics — a synergy that has tremendous potential for teaching us more about how and why organisms evolve.

At first glance, the approach taken by Axelsson and colleagues is simple: to understand what distinguishes domesticated dogs from their wolf ancestors, the authors sequenced the genomes of 12 wolves and 60 dogs and identified DNA fragments that show little variation among dogs, but contain a high density of differences between dogs and wolves (a selective signature). They then determined which genes lie in those fragments.

But the devil is in the details of the dog DNA — the features that underlie a selective signature are exactly the same features that occur after a population bottleneck (the reduction in genetic diversity that occurs when the size of a population is rapidly and drastically reduced). To partition DNA fragments selected for during 10,000 years of domestication from those caused by population bottlenecks during the past several hundred years of breed formation, Axelsson *et al.* created a kind of 'virtual mutt' by choosing 60 dogs that represent 14 diverse breeds and analysing these data as a single large



**Figure 1 | A dog's breakfast.** The winner of the prestigious Westminster Kennel Club dog show traditionally enjoys a celebratory meal at Sardi's Restaurant in New York (the 2012 winner, Malachy, is shown here). Axelsson and colleagues<sup>1</sup> show that dogs are especially well suited to a diet of complex carbohydrates, as a result of selection pressures exerted during thousands of years of domestication.

pool. What should emerge from this analysis are not genomic regions associated with herding, hunting or pointing, but regions that distinguish all dogs from their wild ancestors. The authors identified several functional categories of genes enriched in dog 'candidate domestication regions'. They chose to further investigate those involved in starch metabolism, reasoning that domestication was, in part, a consequence of the agricultural revolution and was therefore facilitated by adaptation to a starch-rich diet (Fig. 1).

Indeed, Axelsson *et al.* report evidence for gain-of-function alterations in dog genes encoding proteins that break down complex carbohydrates (*AMY2B*), hydrolyse oligosaccharides (*MGAM*) and help to transport glucose across the intestinal wall (*SGLT1*). The *AMY2B* result is particularly notable because it seems that the increased activity of the enzyme it encodes, amylase, has occurred as a result of an increase in the number of copies of *AMY2B*. Gene amplification is also known to underlie<sup>4</sup> the increased activity of human amylase in populations in which ancestors consumed diets rich in starch. Thus, the same molecular mechanism has acted on similar genes in different species exposed to the same dietary pressure — a striking example of parallel evolution.

The authors of the *Cell* papers have taken a different tack. They focus on ways to analyse the flood of human sequence data that has resulted from the next-generation sequencing revolution. In cases in which it is possible to determine whether adjacent sequence variants lie on the same chromosome or homologous chromosomes, the range and complexity of potential analytical approaches expand<sup>5,6</sup> and, together with data on the frequency of

specific sequence variants, these approaches can sometimes be used to pinpoint causative variants that underlie recent selective sweeps<sup>7</sup>. (A selective sweep refers to the reduction in genetic variation that occurs in regions adjacent to a mutation that confers a strong selective advantage.) Grossman *et al.*<sup>2</sup> applied this approach to human sequence data from the 1000 Genomes Project to generate a comprehensive catalogue of several hundred regions of around 30 kilobases long, each of which contains dozens of potential adaptive mutations.

The thousands of selected variants that emerge from such analyses highlight the challenge of 'reverse evolutionary genetics' — determining which phenotypic change has been brought about by a specific selected region. Even when causal relationships seem obvious, caution is warranted. For example, the *KITLG* gene, which encodes a crucial signalling molecule for the migration of melanocytes (skin cells that produce the pigment melanin), bears a strong signature of selection in European populations, but was not detected as a skin-colour gene in a recent association study of an African–European admixed population<sup>8</sup>.

Classical genetic studies are the optimal way to establish causal relationships, but in many cases these are impossible because the appropriate populations do not exist. Kamberov *et al.*<sup>3</sup> suggest a way forward using a model organism. Their study focused on a well-recognized mutation in the human gene that encodes the ectodysplasin A receptor (*EDAR*), a signalling molecule that has a role in the development of hair, teeth and exocrine glands<sup>9</sup>. The mutation, which occurred around 30,000 years ago in East Asian populations, results in a valine-to-alanine substitution at

amino-acid-residue 370 of the protein. It has been associated with increased hair thickness<sup>10</sup>, and several lines of evidence suggest that it may have other phenotypic effects<sup>10–12</sup>. Kamberov *et al.* genetically engineered this mutation in mice, and found that the animals had thicker hair, altered mammary-gland morphology and an increased density of eccrine sweat glands compared with normal mice. The authors then closed the phylogenetic circle by showing that the mutation is also associated with an increased number of eccrine sweat glands in a population of Han Chinese. Thus, the phenotypic effect of an adaptive human variant that arose around 30,000 years ago was explored, confirmed and extended to a species whose most recent common ancestor with humans was 3,000 times older than this.

Kamberov and colleagues' study is an exceptional example of experimental genetics, but does it provide, as the authors suggest, a general framework for assessing candidate adaptive mutations? Genetically altered mice are a powerful experimental tool, but the extent to which recent positive selection in humans acts on pathways and amino-acid residues that have been conserved across mammalian evolution is uncertain. More importantly, it is often not clear how to investigate positively selected genomic regions for which the target gene, let alone its action, is unknown. And so a major challenge for population genomics remains the construction of meaningful null hypotheses. As Charles Darwin, the best known evolutionary biologist, once said<sup>13</sup>, "It is always advisable to perceive clearly our ignorance". ■

**Gregory S. Barsh** is at the HudsonAlpha Institute for Biotechnology, Huntsville, 35806 Alabama, and in the Department of Genetics, Stanford University, Stanford, California 94305, USA. **Leif Andersson** is in the Department of Medical Biochemistry and Microbiology, Uppsala University, SE-75123 Uppsala, Sweden.  
e-mails: gbarsh@hudsonalpha.org; leif.andersson@imbim.uu.se

1. Axelsson, E. *et al.* *Nature* **495**, 360–364 (2013).
2. Grossman, S. R. *et al.* *Cell* **152**, 703–713 (2013).
3. Kamberov, Y. G. *et al.* *Cell* **152**, 691–702 (2013).
4. Perry, G. H. *et al.* *Nature Genet.* **39**, 1256–1260 (2007).
5. Sabeti, P. C. *et al.* *Nature* **449**, 913–918 (2007).
6. Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. *PLoS Biol.* **4**, e72 (2006).
7. Grossman, S. R. *et al.* *Science* **327**, 883–886 (2010).
8. Beleza, S. *et al.* *PLoS Genet.* <http://dx.doi.org/10.1371/journal.pgen.1003372> (2013).
9. Mikkola, M. L. *Adv. Exp. Med. Biol.* **691**, 23–33 (2011).
10. Fujimoto, A. *et al.* *Hum. Mol. Genet.* **17**, 835–843 (2008).
11. Mou, C. *et al.* *Hum. Mutat.* **29**, 1405–1411 (2008).
12. Chang, S. H., Jobling, S., Brennan, K. & Headon, D. J. *PLoS ONE* **4**, e7591 (2009).
13. Darwin, C. *The Expression of the Emotions in Man and Animals* (Murray, 1872).



# Functional organization of human sensorimotor cortex for speech articulation

Kristofer E. Bouchard<sup>1,2</sup>, Nima Mesgarani<sup>1,2</sup>, Keith Johnson<sup>3</sup> & Edward F. Chang<sup>1,2,4</sup>

**Speaking is one of the most complex actions that we perform, but nearly all of us learn to do it effortlessly. Production of fluent speech requires the precise, coordinated movement of multiple articulators (for example, the lips, jaw, tongue and larynx) over rapid time scales. Here we used high-resolution, multi-electrode cortical recordings during the production of consonant-vowel syllables to determine the organization of speech sensorimotor cortex in humans. We found speech-articulator representations that are arranged somatotopically on ventral pre- and post-central gyri, and that partially overlap at individual electrodes. These representations were coordinated temporally as sequences during syllable production. Spatial patterns of cortical activity showed an emergent, population-level representation, which was organized by phonetic features. Over tens of milliseconds, the spatial patterns transitioned between distinct representations for different consonants and vowels. These results reveal the dynamic organization of speech sensorimotor cortex during the generation of multi-articulator movements that underlies our ability to speak.**

Speech communication critically depends on the ability to produce the large number of sounds that compose a given language<sup>1,2</sup>. The wide range of spoken sounds results from highly flexible configurations of the vocal tract, which filters sound produced at the larynx through movements of the lips, jaw and tongue that are coordinated precisely<sup>3–5</sup>. Each articulator has extensive degrees of freedom, making a large number of different speech movements possible. How humans exert such precise control despite the wide variety of movement possibilities is a central unanswered question<sup>1,6,7</sup>.

The cortical control of articulation is mediated primarily by the ventral half of the lateral sensorimotor (Rolandic) cortex (ventral sensorimotor cortex, vSMC)<sup>8–10</sup>, which provides corticobulbar projections to, and afferent innervation from, the face and vocal tract (Fig. 1a, b)<sup>11,12</sup>. The U-shaped vSMC is composed of the pre- and post-central gyri (Brodmann areas 1, 2, 3 and 6b), and the gyrus directly ventral to the termination of the central sulcus called the *gyrus* (Brodmann area 43) (Fig. 1a, b)<sup>13</sup>. Using electrical stimulation, Foerster and Penfield described the somatotopic organization of face and mouth representations in human vSMC<sup>14,15,16</sup>. However, focal stimulation could not evoke meaningful utterances, implying that speech is not stored in discrete cortical areas. Instead, the production of phonemes and syllables is thought to arise from a coordinated motor pattern involving multiple articulator representations<sup>1,3,4,5,9</sup>.

To understand the functional organization of vSMC in articulatory sensorimotor control, we recorded neural activity directly from the cortical surface in three human subjects implanted with high-density multi-electrode arrays as part of their preparation for epilepsy surgery (Fig. 1a). Intracranial cortical recordings were synchronized with microphone recordings as subjects read aloud consonant-vowel syllables (19 consonants followed by /a/, /u/ or /i/; Supplementary Fig. 1) that are commonly used in American English. This task was designed to sample across a range of phonetic features, including different constriction locations (place of articulation) and different constriction degrees or shapes (manner of articulation) for a given articulatory organ<sup>17,18,19</sup>.

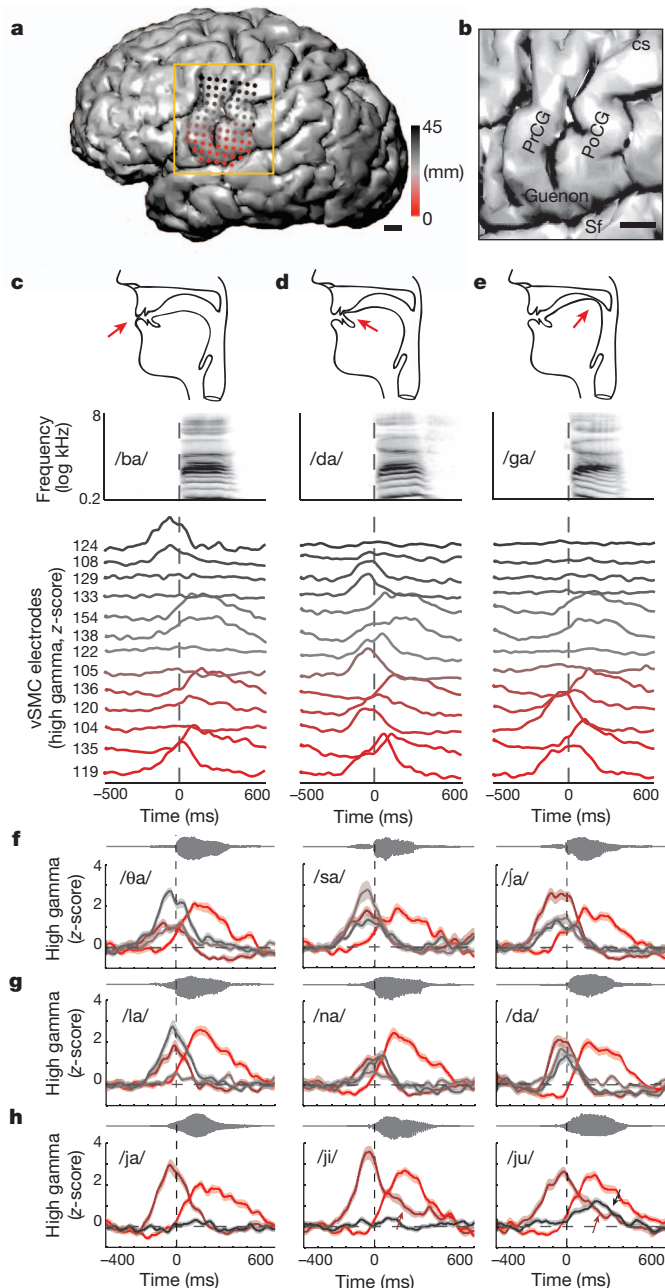
## vSMC physiology during syllable production

We aligned cortical recordings to acoustic onsets of consonant-to-vowel transitions ( $t = 0$ ) to provide a common reference point across consonant-vowel syllables (Fig. 1c–e). We focused on the high-gamma frequency component of local field potentials (85–175 Hz)<sup>20,21,22</sup>, which correlates well with multi-unit firing rates<sup>23</sup>. For each electrode, we normalized the time-varying high-gamma amplitude to baseline statistics by transforming to z-scores.

During syllable articulation, approximately 30 active vSMC electrode sites were identified per subject (approximately 1,200 mm<sup>2</sup>, change in z-score of greater than 2 for any syllable). Cortical activity from selected electrodes distributed along the vSMC dorsoventral axis is shown for /ba/, /da/ and /ga/ (Fig. 1c–e, same colouring as in Fig. 1a). The plosive consonants (/b/, /d/, /g/) are produced by transient occlusion of the vocal tract by the lips, front tongue and back tongue, respectively, whereas the vowel /a/ is produced by a low, back tongue position during phonation. Dorsally located electrodes (for example, Fig. 1c–e, electrodes 124 and 108; black) were active during production of /b/, which requires transient closure of the lips. In contrast, mid-positioned electrodes (for example, electrodes 129, 133 and 105; grey) were active during production of /d/, which requires forward tongue protrusion against the alveolar ridge. A more ventral electrode (for example, electrode 104; red) was most active during production of /g/, which requires a posterior-oriented tongue elevation towards the soft palate. Other electrodes appear to be active during the vowel phase for /a/ (for example, electrodes 154, 136 and 119).

Cortical activity at different electrode subsets was superimposed to visualize spatiotemporal patterns across other phonetic contrasts. Consonants produced with different constriction locations of the tongue tip, (for example, /θ/ (dental), /s/ (alveolar), and /j/ (post-alveolar)), showed specificity across different electrodes in central vSMC (Fig. 1f), although they were not as categorical as those shown for consonants involving different articulators in Fig. 1c–e. Consonants with similar tongue constriction locations, but different constriction degree or constriction shape, were generated by overlapping electrode sets exhibiting

<sup>1</sup>Department of Neurological Surgery and Department of Physiology, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, California 94143, USA. <sup>2</sup>Center for Integrative Neuroscience, 675 Nelson Rising Lane, University of California, San Francisco, California 94158, USA. <sup>3</sup>Department of Linguistics, University of California, Berkeley, 1203 Dwinelle Hall, Berkeley, California 94720, USA. <sup>4</sup>UCSF Epilepsy Center, University of California, San Francisco, 400 Parnassus Avenue, San Francisco, California 94143, USA.



**Figure 1 | vSMC physiology during syllable production.** **a**, Magnetic resonance imaging (MRI) reconstruction of a single subject brain with vSMC electrodes (dots), coloured according to distance from the Sylvian fissure (black and red are the most dorsal and ventral positions, respectively). **b**, Expanded view of vSMC anatomy. cs, central sulcus; PoCG, post-central gyrus; PrCG, pre-central gyrus; Sf, Sylvian fissure. Scale bars, 1 cm. **c–e**, Top, vocal tract schematics for three consonants (/b/, /d/, /g/), produced by occlusion at the lips, tongue tip and tongue body, respectively (red arrow). Middle, spectrograms of spoken consonant-vowel syllables. Bottom, average cortical activity from a subset of electrodes (electrode number on far right, same colouring as in **a**). Vertical dashed line, acoustic onset of consonant-vowel transition. **f–h**, Cortical activity at selected electrodes for different phonetic contrasts (mean  $\pm$  s.e.m.). Acoustic waveforms are displayed above. **f**, Fricatives (/θ/ ('th' of 'thin'), /s/, /ʃ/ ('sh' of 'shin')) with different constriction locations. **g**, Front tongue consonants (/l/, /n/, /d/) with different constriction degree or shapes. **h**, Single consonant (/j/ ('y' of 'yes')) with different vowels (/a/, /i/, /u/). Purple arrows correspond to a tongue electrode with prolonged activity for /i/ and /u/ vowels. Black arrow corresponds to an active lip electrode for /u/.

different relative activity magnitudes (Fig. 1g, /l/ (lateral) versus /n/ (nasal stop) versus /d/ (oral stop)). Syllables with the same consonant followed by different vowels (Fig. 1h, /ja/, /ji/, /ju/) were found to have similar activity patterns before the consonant-vowel transition. During vowel phonation, a dorsal electrode is clearly active during /u/, but not /i/ or /a/ (Fig. 1h, /ju/; black arrow) whereas another electrode in the middle of vSMC had prolonged activity during /i/ and /u/ vowels compared to /a/ (Fig. 1h, /ji/ and /ju/; purple arrows). These contrasting examples show that important phonetic properties can be observed qualitatively from the rich repertoire of vSMC spatiotemporal patterns.

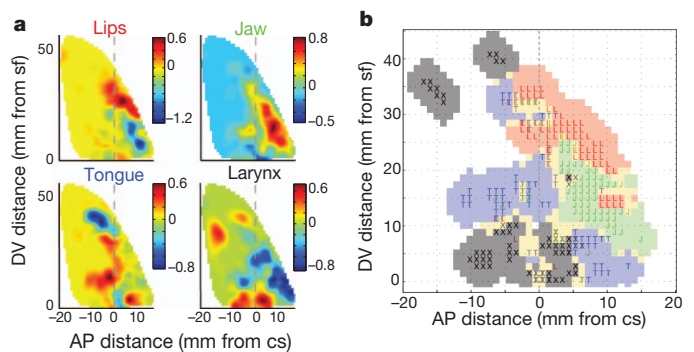
## Spatial representation of articulators

To determine the spatial organization of speech-articulator representations, we examined how cortical activity at each electrode depended on the movement of a given articulator (using a general linear model). We assigned binary variables to four articulatory organs (lips, tongue, larynx and jaw) that are used in producing the consonant component of each consonant-vowel syllable (Supplementary Fig. 1). The spatial distribution of optimal weightings for these articulators (averaged over time and subjects) were plotted as a function of dorsoventral distance from the Sylvian fissure and anteroposterior distance from the central sulcus. We found representations for each articulator distributed across vSMC (Fig. 2a). For example, the lip representation was localized to the dorsal aspect of vSMC, whereas the tongue representation was distributed more broadly than the lip representation across the ventral aspect.

To determine topographic organization of articulators across subjects, we extracted the greatest 10% of weightings from individual articulator distributions (Fig. 2a) and used a clustering algorithm (*k*-nearest neighbour) to classify the surrounding cortex (Fig. 2b). We found an overall somatotopic dorsoventral arrangement of articulator representations laid out in the following sequence: larynx, lips, jaw, tongue and larynx (Fig. 2a, b and Supplementary Figs 2–5). An analysis of the fractional representation of all articulators at single electrodes showed a clear tuning preference for individual articulators at single electrodes and also demonstrated that single electrodes had functional representations of multiple articulators (Supplementary Fig. 6).

## Timing of articulator representations

As the time course of articulator movements is on the scale of tens of milliseconds, previous approaches have been unable to resolve temporal properties associated with individual articulator representations. We examined the timing of correlations between cortical activity



**Figure 2 | Spatial representation of articulators.** **a**, Localization of lips, jaw, tongue and larynx representations. Average magnitude of articulator weightings (colour scale) plotted as a function of anteroposterior (AP) distance from the central sulcus and dorsoventral (DV) distance from the Sylvian fissure ( $n = 3$  subjects). **b**, Functional somatotopic organization of speech-articulator representations in vSMC. Lips (L, red); jaw (J, green); tongue (T, blue); larynx (X, black); mixed (yellow). Letters correspond to locations, based on direct measurement-derived regression weights; shaded rectangles correspond to regions classified by *k*-nearest neighbour.

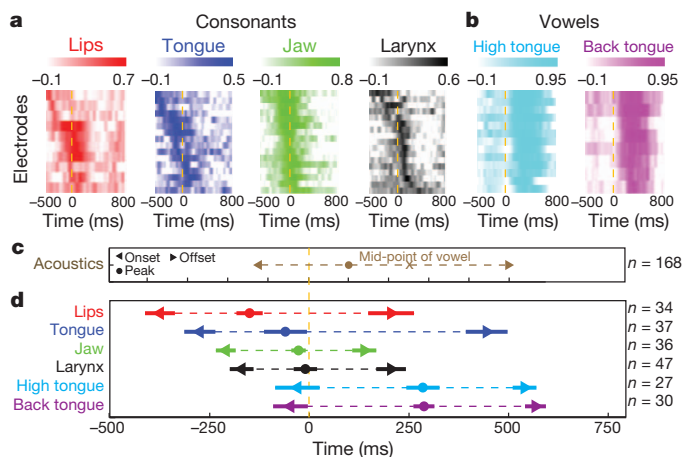


and specific consonant articulators (using partial correlation analysis), and included two vowel articulatory features (back tongue and high tongue; Supplementary Fig. 1).

Time courses of correlations were plotted for electrodes with highest values, sorted by onset latency (Fig. 3a). We found that jaw, high tongue and back tongue had very consistent timing across electrodes. Similar results were found for tongue, lips and larynx, but with more variable latencies. Timing relationships between articulator representations were staggered, reflecting a temporal organization during syllable production: lip and tongue correlations began well before sound onset (Fig. 3a, c, d); jaw and larynx correlations were aligned to the consonant-vowel transition (Fig. 3a, c, d); and high tongue and back tongue features showed high temporal specificity for the vowel phase, peaking near the acoustic mid-point of the vowels (approximately 250 ms, Fig. 3b–d). This sequence of articulator correlations was consistent across subjects (Fig. 3d,  $P < 10^{-10}$ , analysis of variance (ANOVA),  $F = 40$ , d.f. = 5,  $n = 211$  electrodes from 3 subjects) and is in accordance with the timing of articulator movements shown in speech-kinematics studies<sup>3,5,17,24</sup>. We found no statistically significant onset-latency differences in those areas 10 mm anterior and posterior to the central sulcus or across the guenon ( $P > 0.4$ , rank-sum test;  $n = 71$  and  $n = 67$ , respectively; Supplementary Fig. 7). This is consistent with mixed sensory and motor orofacial responses throughout vSMC, which are also seen in stimulation experiments<sup>14,25</sup>.

### Phonetic organization of spatial patterns

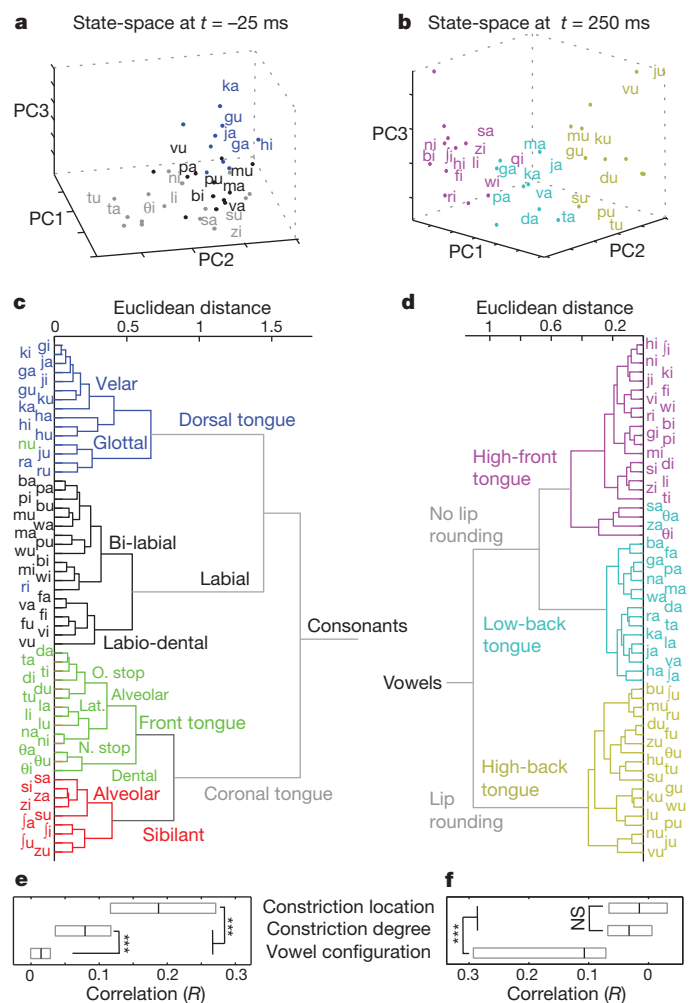
The distributed organization of speech articulator representations (Fig. 2) led us to propose that coordination of the multiple articulators required for speech production would be associated with spatial patterns of cortical activity. We refer here to this population-derived pattern as the phonetic representation. To determine its organizational properties, we used principal component analysis to transform the observed cortical activity patterns into a ‘cortical state-space’ (approximately 60% of variance is explained by 9 spatial principal components for all subjects, Supplementary Figs 8 and 9)<sup>26–30</sup>. *k*-means clustering during the consonant phase (25 ms before the consonant-vowel transition,  $t = -25$  ms) showed that the cortical state-space was organized into three clusters (quantified by silhouette analysis) corresponding to the major oral articulators: labial, coronal tongue, and dorsal tongue (Fig. 4a and Supplementary Fig. 10).



**Figure 3 | Temporal representation of articulators.** **a, b**, Timing of correlations between cortical activity and consonant (**a**) and vowel (**b**) articulator features. Colour maps display correlation coefficients ( $R$ ) for a subset of electrodes. **c**, Acoustic landmarks. Onset (end of arrows, left), peak power (shown by a dot in each case) and offset (end of arrows, right) for consonant-vowel syllables (mean  $\pm$  s.e.m.,  $n = 168$  syllables, all subjects). Error bars are smaller than the symbols. **d**, Temporal sequence and range of correlations. Symbols are as in **c**. Data are mean (symbols)  $\pm$  s.e.m. (thick solid line) across electrodes from all subjects.

During the vowel phase (250 ms after the consonant-vowel transition,  $t = 250$ ), we found clear separation of /a/, /i/ and /u/ vowel states (Fig. 4b). Similar clustering of consonants and vowels was found across subjects ( $P < 10^{-10}$  for clustering of both consonants and vowels, Supplementary Fig. 11).

Theories of speech motor control and phonology have speculated that there is a hierarchical organization of phoneme representations, given the anatomical and functional dependencies of the vocal tract articulators during speech production<sup>3,4,17,18,31</sup>. To evaluate such organization in vSMC, we applied hierarchical clustering to the cortical state-space (Fig. 4c, d). For consonants, this analysis confirmed that the primary tier of organization was defined by the major oral articulator features: dorsal, labial or coronal (Fig. 4c). These major articulators were superordinate to the constriction location within each articulator. For example, the labial cluster could be subdivided



**Figure 4 | Phonetic organization of spatial patterns.** **a, b**, Scatterplots of consonant-vowel syllables in the first three principal components for consonants (25 ms before consonant-vowel transition) (**a**) and vowels (250 ms after consonant-vowel transition) (**b**). A subset of consonant-vowels are labelled with international phonetic alphabet (IPA) symbols, all others have dots. Colouring denotes *k*-means cluster membership. **c, d**, Hierarchical clustering of cortical state-space at consonant (25 ms before consonant-vowel transition) and vowel time points (250 ms after consonant-vowel transition). Individual syllables and dendrogram branches are colour-coded and labelled by known linguistic categories using the same colour scheme as in Fig 4a, b, with new subdivisions of the coronal tongue into front tongue and sibilant (in green and red, respectively). Lat, lateral; N. stop, nasal stop; O. stop, oral stop. **e, f**, Correlations between cortical state-space and phonetic features. Black vertical lines, medians; grey boxes, 25th and 75th percentiles. \*\*\* $P < 10^{-10}$ , WSRT;  $n = 297$  for both consonants and vowels. NS, not significant.

into bi-labial and labiodental. Only at the lowest level of the hierarchy did we observe suggestions of organization according to constriction degree or shape, such as the sorting of nasal (/n/ syllables), oral stops (/d/, /t/) and lateral approximants (/l/). Similarly, during the vowel period, a primary distinction was based on the presence or absence of lip rounding (/u/ versus /a/ and /i/), and a secondary distinction was based on tongue posture (height, and front or back position) (Fig. 4d). Therefore, the major oral articulator features that organize consonant representations are similar to those for vowels.

Across an early time period (375 ms before, to 120 ms after, the consonant-vowel transition), we found that consonant features describing constriction location had a significantly greater correlation with the cortical state-space than constriction degree, which in turn was significantly more correlated than the upcoming vowel ( $P < 10^{-10}$ , Wilcoxon signed-rank test (WSRT),  $n = 297$  from 3 subjects; see Supplementary Fig. 12 for phonetic feature sets). This analysis shows that constriction location accounts for more of the structure of spatial activity patterns than does constriction degree or shape. Similarly, across a later time period (125 ms to 620 ms after the consonant-vowel transition), we found that vowel features provided the greatest correlation (vowel configuration versus other feature sets,  $P < 10^{-10}$ , WSRT,  $n = 297$  from 3 subjects).

### Dynamics of phonetic representations

The dynamics of neural populations have provided insights into the structure and function of many neural circuits<sup>6,26,27,29,32,33</sup>. To determine the dynamics of phonetic representations, we investigated how state-space trajectories for consonants and vowels entered and departed target regions for phonetic clusters. Trajectories of individual consonant-vowel syllables were visualized by plotting their locations in the first two principal-component dimensions versus time (Fig. 5a, b; principal component 1 (PC1) and PC2 for one of the subjects).

We examined first how trajectories of different consonants transitioned to a single vowel, /u/ (Fig. 5a). The cortical state-space was initially unstructured, and then individual trajectories converged within phonetic clusters (for example, labial, front tongue, dorsal tongue and sibilant), and at the same time trajectories for different clusters diverged from one another. These convergent and divergent dynamics gradually increased the separability of different phonetic clusters (the mean difference of between-cluster and within-cluster distances). Later, as each consonant transitioned to /u/, trajectories

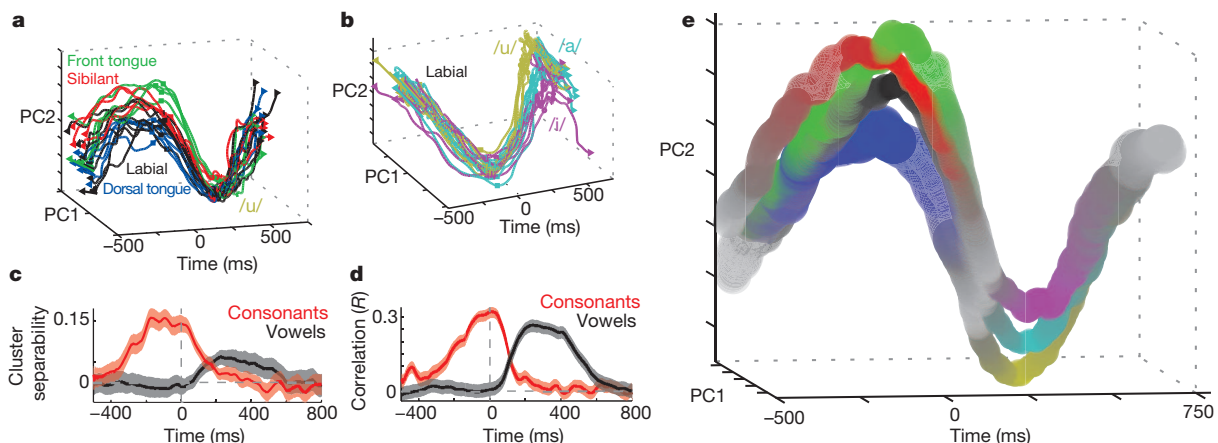
converged to a compact target region for the vowel. Finally, trajectories diverged randomly, presumably as articulators returned to neutral position. Analogous dynamics were observed during the production of a single consonant cluster (for example, labials) transitioning to different vowels (/a/, /i/ and /u/) (Fig. 5b).

We quantified the internal dynamical properties of the cortical state-space by calculating cluster separability. The time course of cluster separability, averaged across subjects and consonant-vowel syllables (Fig. 5c) showed that separability peaked approximately 200 ms before the consonant-vowel transition for consonants (onset, approximately 300 ms before the consonant-vowel transition), and at 250 ms after the consonant-vowel transition for vowels (onset, approximately 50 ms after the consonant-vowel transition). We examined further the dynamics of correlations between the structure of the cortical state-space and phonetic features (averaged across subjects) (plotted in Fig. 5d). Across subjects, we found that cluster separability and the correlation between cortical state-space organization and phonetic features were tightly linked for both consonants and vowels in a time-dependent fashion ( $R^2$  range = 0.42–0.98,  $P < 10^{-10}$  in all cases). This shows that the dynamics of clustering in the cortical state-space is coupled strongly to the degree to which the cortical state reflects the phonetic structure of the vocalization.

Visualization of the dynamic structure of the cortical state-space during production of all consonant-vowel syllables (Fig. 5e) showed that, as the cortical state comes to reflect phonetic structure, different phonetic clusters diverge from one another, while the trajectories within the clusters converge. Furthermore, we observed correlates of the earlier articulatory specification for sibilants (/j/, /z/, /s/). In addition, with all consonant-vowel syllables on the same axes, we observed that in comparison to vowels, consonants occupy a distinct region of cortical state-space, despite sharing the same articulators. The distribution of state-space distances was significantly greater in consonant-vowel comparisons than in consonant-consonant or vowel-vowel comparisons ( $P < 10^{-10}$  for all comparisons, WSRT,  $n = 4623$  in all cases, Supplementary Fig. 11). Finally, the consonant-to-vowel sequence reveals a periodic structure, which is sub-specified for consonant and vowel features.

### Discussion

Our broad-coverage, high-resolution direct cortical recordings enabled us to examine the spatial and temporal profiles of speech articulator representations in human vSMC. Cortical representations are



**Figure 5 | Dynamics of phonetic representations.** **a, b,** Cortical state-space trajectories. **a,** Consonants transitioning to the vowel /u/. Each line corresponds to a single consonant-vowel trajectory. For each line, the left triangle indicates  $t = -500$  ms, the square indicates  $t = -25$  ms, the circle indicates  $t = 250$  ms, and the right triangle indicates  $t = 750$  ms. **b,** Trajectories of the labial consonants transitioning to /a/, /i/ and /u/. **c, d,** Across-subject averages of cluster separability (**c**) and correlation between cortical state-space structure

and phonetic features (**d**) for consonants and vowels (mean  $\pm$  s.e.m.). **e,** Time-course of consonant-vowel syllable trajectories for one subject. Each colour corresponds to one of the consonant or vowel groups (colours are the same as in **a** and **b** above). The centre of each coloured tube is located at the centroid of the corresponding phonetic cluster. Tube diameter corresponds to cluster density and colour saturation represents the correlation between the structure of the cortical state-space and phonetic features.



somatotopically organized, with individual sites tuned for a preferred articulator and co-modulated by other articulators. The dorsoventral layout of articulator representations recapitulates the rostral-to-caudal layout of the vocal tract. However, we found an additional laryngeal representation located at the dorsal-most end of vSMC<sup>8,10,34,35</sup>. This dorsal laryngeal representation seems to be absent in non-human primates<sup>11,36,37</sup>, suggesting a unique feature of human vSMC for the specialized control of speech. Pre- and post-central gyrus neural activity occurred before vocalization, which may reflect the integration of motor commands with proprioceptive information for rapid feedback control during speaking<sup>9,38–43</sup>.

Just as focal stimulation is insufficient to evoke speech sounds, it is not any single articulator representation, but the coordination of multiple articulator representations across the vSMC network that generates speech. Analysis of spatial patterns of activity showed an emergent hierarchy of network states that organizes phonemes by articulatory features. This functional hierarchy of network states contrasts with the anatomical hierarchy often considered in motor control<sup>44</sup>. The cortical state-space organization probably reflects the coordinative patterns of articulatory motions during speech, and is notably similar to a theorized cross-linguistic hierarchy of phonetic features ('feature geometry')<sup>3,18,31,45</sup>. In particular, the findings support gestural theories of speech control<sup>3</sup> over alternative acoustic (a hierarchy organized primarily by constriction degree)<sup>19</sup> or vocal-tract geometry theories (no hierarchy of constriction location and degree)<sup>18</sup>.

The vSMC population showed convergent and divergent dynamics during the production of different phonetic features. The dynamics of individual phonemes were superimposed on a slower oscillation that characterizes the transition between consonants and vowels. Although trajectories were found to originate or terminate in different regions, they consistently pass through the same (target) region of the state-space for shared phonetic features<sup>46</sup>. Consonants and vowels occupy distinct regions of the cortical state-space. Large state-space distances between consonant and vowel representations may explain why it is more common in speech errors to substitute consonants with one another, and vowels with vowels, but very rarely consonants with vowels or vowels with consonants (that is, in 'slips of the tongue')<sup>47</sup>.

We have shown that a relatively small set of articulator representations can combine flexibly to create the large variety of speech sounds in American English. The major organizational features found here define phonologies of languages from across the world<sup>31</sup>. Consequently, these cortical organizational principles are likely to be conserved, with further specification for unique articulatory properties across different languages.

## METHODS SUMMARY

Three subjects underwent surgical placement of subdural arrays as part of their clinical treatment for epilepsy (see Supplementary Table 1 for clinical details). Statistical tests were considered significant if the Bonferroni corrected rate of incorrectly rejecting the null hypothesis was less than 0.05.

**Full Methods** and any associated references are available in the online version of the paper.

Received 23 June 2012; accepted 15 January 2013.

Published online 20 February 2013.

- Levelt, W. J. M. *Speaking: From Intention to Articulation* (MIT Press, 1993).
- Ladefoged, P. & Johnson, K. *A Course in Phonetics* (Wadsworth Publishing, 2010).
- Browman, C. P. & Goldstein, L. Articulatory gestures as phonological units. *Haskins Laboratories Status Report on Speech Research* **99**, 69–101 (1989).
- Fowler, C. A., Rubin, P. E., Remez, R. E. & Turvey, M. T. in *Language Production: Speech and Talk* Vol. 1 (ed. Butterworth, B.) 373–420 (Academic Press, 1980).
- Gracco, V. L. & Lofqvist, A. Speech motor coordination and control: evidence from lip, jaw, and laryngeal movements. *J. Neurosci.* **14**, 6585–6597 (1994).
- Schöner, G. & Kelso, J. A. Dynamic pattern generation in behavioral and neural systems. *Science* **239**, 1513–1520 (1988).
- Franklin, D. W. & Wolpert, D. M. Computational mechanisms of sensorimotor control. *Neuron* **72**, 425–442 (2011).
- Brown, S. et al. The somatotopy of speech: phonation and articulation in the human motor cortex. *Brain Cogn.* **70**, 31–41 (2009).
- Guenther, F. H., Ghosh, S. S. & Tourville, J. A. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* **96**, 280–301 (2006).
- Schulz, G. M., Varga, M., Jeffries, K., Ludlow, C. L. & Braun, A. R. Functional neuroanatomy of human vocalization: an H2150 PET study. *Cereb. Cortex* **15**, 1835–1847 (2005).
- Jürgens, U. Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.* **26**, 235–258 (2002).
- Kuyper, H. G. Corticobulbar connexions to the pons and lower brain-stem in man: an anatomical study. *Brain* **81**, 364–388 (1958).
- Brodman, K. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues* (Smith-Gordon, 1994).
- Penfield, W. & Boldrey, E. Somatic motor and sensory representation in the cerebral cortex of man studied by electrical stimulation. *Brain* **60**, 389–443 (1937).
- Foerster, O. The cerebral cortex in man. *Lancet* **221**, 309–312 (1931).
- Penfield, W. & Roberts, R. *Speech and Brain: Mechanisms*. (Princeton, 1959).
- Saltzman, E. & Munhall, K. A dynamical approach to gestural patterning in speech production. *Ecol. Psychol.* **1**, 333–382 (1989).
- Clements, G. N. & Hume, E. in *The Handbook of Phonological Theory* (ed. Goldsmith, J. A.) 245–306 (Basil Blackwell, 1995).
- Chomsky, N. & Halle, M. *The Sound Pattern of English* (MIT Press, 1991).
- Mesgarani, N. & Chang, E. F. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* **485**, 233–236 (2012).
- Crone, N. E., Miglioretti, D. L., Gordon, B. & Lesser, R. P. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band. *Brain* **121**, 2301–2315 (1998).
- Edwards, E. et al. Spatiotemporal imaging of cortical activation during verb generation and picture naming. *Neuroimage* **50**, 291–301 (2010).
- Ray, S. & Maunsell, J. H. Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol.* **9**, e1000610 (2011).
- Kent, R. D. in *The Production of Speech* (ed. MacNeilage, P. F.) (Springer-Verlag, 1983).
- McCarthy, G., Allison, T. & Spencer, D. D. Localization of the face area of human sensorimotor cortex by intracranial recording of somatosensory evoked potentials. *J. Neurosurg.* **79**, 874–884 (1993).
- Afshar, A. et al. Single-trial neural correlates of arm movement preparation. *Neuron* **71**, 555–564 (2011).
- Mazor, O. & Laurent, G. Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons. *Neuron* **48**, 661–673 (2005).
- Sussillo, D. & Abbott, L. F. Generating coherent patterns of activity from chaotic neural networks. *Neuron* **63**, 544–557 (2009).
- Ahrens, M. B. et al. Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature* **485**, 471–477 (2012).
- Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Ryu, S. I. & Shenoy, K. V. Cortical preparatory activity: representation of movement or first cog in a dynamical machine? *Neuron* **68**, 387–400 (2010).
- McCarthy, J. J. Feature geometry and dependency: a review. *Phonetica* **45**, 84–108 (1988).
- Briggman, K. L. & Kristan, W. B. Multifunctional pattern-generating circuits. *Annu. Rev. Neurosci.* **31**, 271–294 (2008).
- Churchland, M. M. et al. Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
- Brown, S., Ngan, E. & Liotti, M. A larynx area in the human motor cortex. *Cereb. Cortex* **18**, 837–845 (2008).
- Terumitsu, M., Fujii, Y., Suzuki, K., Kwee, I. L. & Nakada, T. Human primary motor cortex shows hemispheric specialization for speech. *Neuroreport* **17**, 1091–1095 (2006).
- Hast, M. H., Fischer, J. M., Wetzel, A. B. & Thompson, V. E. Cortical motor representation of the laryngeal muscles in Macaca mulatta. *Brain Res.* **73**, 229–240 (1974).
- Jürgens, U. On the elicibility of vocalization from the cortical larynx area. *Brain Res.* **81**, 564–566 (1974).
- Pruszynski, J. A. et al. Primary motor cortex underlies multi-joint integration for fast feedback control. *Nature* **478**, 387–390 (2011).
- Hatsopoulos, N. G. & Suminski, A. J. Sensing with the motor cortex. *Neuron* **72**, 477–487 (2011).
- Tremblay, S., Shiller, D. M. & Ostry, D. J. Somatosensory basis of speech production. *Nature* **423**, 866–869 (2003).
- Matyas, F. et al. Motor control by sensory cortex. *Science* **330**, 1240–1243 (2010).
- Rathelot, J. A. & Strick, P. L. Muscle representation in the macaque motor cortex: an anatomical perspective. *Proc. Natl Acad. Sci. USA* **103**, 8257–8262 (2006).
- Gracco, V. L. & Abbs, J. H. Dynamic control of the perioral system during speech: kinematic analyses of autogenic and nonautogenic sensorimotor processes. *J. Neurophysiol.* **54**, 418–432 (1985).
- Sherrington, C. S. *The Integrative Action of the Nervous System* (Yale University Press, 1911).
- Jakobson, R., Fant, G. & Halle, M. *Preliminaries to speech analysis: the distinctive features and their correlates* (MIT Press, 1969).
- Keating, P. A. *The Window Model of Coarticulation: Articulatory Evidence* (Cambridge Univ. Press, 1990).
- Dell, G. S., Juliano, C. & Govindjee, A. Structure and content in language production: a theory of frame constraints in phonological speech errors. *Cogn. Sci.* **17**, 149–195 (1993).

Supplementary Information is available in the online version of the paper.

**Acknowledgements** We thank A. Ren for technical help with data collection and pre-processing, and M. Babiak for audio transcription. J. Houde, C. Niziolek, S. Lisberger, K. Chaisanguanthum, C. Cheung and I. Garner provided helpful comments on the manuscript. E.F.C. was funded by the US National Institutes of Health grants R00-NS065120, DP2-OD00862 and R01-DC012379, and the Ester A. and Joseph Klingenstein Foundation.

**Author Contributions** E.F.C. conceived and collected the data for this project. K.E.B. designed and implemented the analysis with assistance from E.F.C. N.M. assisted with

preliminary analysis. K.E.B. and E.F.C. wrote the manuscript. K.J. provided phonetic consultation on experimental design and interpretation of results. E.F.C. supervised the project.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.F.C. (ChangEd@neurosurg.ucsf.edu).



## METHODS

The experimental protocol was approved by the Human Research Protection Program at the University of California, San Francisco.

**Subjects and experimental task.** Three native-English-speaking human subjects underwent chronic implantation of a high-density, subdural electrocorticographic (ECoG) array over the left hemisphere (two subjects) or right hemisphere (one subject) as part of their clinical treatment of epilepsy (see Supplementary Table 1 for clinical details)<sup>48</sup>. Subjects gave their written informed consent before the day of surgery. All subjects had self-reported normal hearing and underwent neuropsychological language testing (including the Boston naming and verbal fluency tests) and scored within the range considered normal. Each subject read aloud consonant-vowel syllables composed of 18 or 19 consonants (19 consonants for two subjects, 18 consonants for one subject), followed by one of three vowels. Each consonant-vowel syllable was produced between 15 and 100 times. Microphone recordings were synchronized with the multi-channel ECoG data.

**Data acquisition and signal processing.** Cortical local field potentials (LFPs) were recorded with ECoG arrays and a multi-channel amplifier connected optically to a digital signal processor (Tucker-Davis Technologies). The spoken syllables were recorded with a microphone, amplified digitally, and recorded simultaneously with the ECoG data. ECoG signals were acquired at 3,052 Hz.

The time series from each channel was inspected visually and quantitatively for artefacts or excessive noise (typically 60 Hz line noise). These channels were excluded from all subsequent analysis and the raw recorded ECoG signal of the remaining channels were then common-average referenced and used for spectro-temporal analysis. For each (useable) channel, the time-varying analytic amplitude was extracted from eight bandpass filters (Gaussian filters, logarithmically increasing centre frequencies (85–175 Hz) and semi-logarithmically increasing bandwidths) with the Hilbert transform. The high-gamma (high- $\gamma$ ) power was then calculated by averaging the analytic amplitude across these eight bands, and then this signal was down-sampled to 200 Hz. High- $\gamma$  power was z-scored relative to the mean and standard deviation of baseline data for each channel. Throughout the Methods, high- $\gamma$  power refers to this z-scored measure.

**Acoustic analysis.** The recorded speech signal was transcribed off-line using WaveSurfer (<http://www.speech.kth.se/wavesurfer/>). The onset of the consonant-to-vowel transition was used as the common temporal reference point for all subsequent analysis (see Supplementary Fig. 1). This was chosen because it permits alignment across all of the syllables and allows for a consistent discrimination of the consonantal and vocalic components. Post-hoc analysis of acoustic timing revealed the onset of the consonant-to-vowel transition to be highly reproducible across multiple renditions of the same syllable. As such, alignment at the consonant-to-vowel transition results in relatively small amounts of inter-syllable jitter in estimated times of acoustic onset, offset and peak power.

For temporal analysis of the consonant-vowel acoustic structure, each individual vocalization was first converted to a cochlear spectrogram by passing the sound-pressure waveform through a filter bank emulating the cochlear transfer function<sup>49</sup>. As the current analysis of cortical data leverages the cross-syllabic variability in (average) high- $\gamma$  (see below), we reduced the data set of produced vocalizations to a single exemplar for each consonant-vowel syllable. For each unique consonant-vowel syllable, the cochlear spectrograms associated with each utterance of that consonant-vowel ( $S_i(t,f)$ ) were analysed to find a single prototypical example (Pspct), defined as the syllable that had the minimum (min) spectrotemporal difference from every other syllable of that kind:

$$\text{Pspct} = \min_i \left( \sum_{j,f} (S_j(t,f) - S_i(t,f))^2 \right) \quad (1)$$

where,  $S_i(t,f)$  is the spectrogram of the  $i$ th example of the syllable, corresponding to the power at time  $t$  and frequency  $f$ . The onset, peak and offset of acoustic power were extracted for each syllable prototype using a thresholding procedure.

**Articulator state matrix and phonetic feature matrix.** To describe the engagement of the articulators in the production of different consonant-vowel syllables, we drew from standard descriptions of the individual consonant and vowel sounds in the International Phonetic Alphabet (IPA)<sup>50</sup>. Each consonant-vowel syllable was associated with a binary vector describing the engagement of the speech articulators used to produce the consonant-vowel syllable. For the linear analyses presented in Figs 2 and 3, the articulator state vector ( $B_i$ ) for each consonant-vowel syllable  $s_i$  was defined by six binary variables describing the four main articulator organs (lips, tongue, larynx and jaw) for consonant production and two vocalic tongue configurations (high tongue and back tongue) (Supplementary Fig. 1). Although more detailed descriptions are possible (for example, alveolar-dental), the linear methods used for these analyses require the articulator variables to be linearly independent (no feature can be described as an exact linear combination of the others), although the features may have correlations. An expanded phonetic

feature matrix (nine consonant constriction location variables, six consonant constriction degree or shape variables, and four vowel tongue and lip configuration variables; derived from the IPA, Supplementary Fig. 12) was used in the non-parametric analysis of the cortical state-space (Figs 4 and 5).

**Spatial organization derived from a general linear model.** To examine the spatial organization with which high- $\gamma$  was modulated by the engagement of the articulators, we determined how the activity of each electrode varied with consonant articulator variables using a general linear model (GLM). Here, at each moment in time ( $t$ ), the GLM described the high- $\gamma$  of each electrode as an optimally weighted sum of the articulators engaged during speech production. High- $\gamma(t)$  ( $H\gamma(t)$ ) recorded on each electrode ( $e_i$ ), during the production of syllable  $s_j$ ,  $H\gamma_{i,j}(t)$ , was modelled as a linear weighted sum of the binary vector associated with the consonant component of  $s_j$ , ( $B_j^c$ ):

$$H\gamma_{i,j}(t) = \beta_i(t) \bullet B_j^c + \beta_{i0}(t) \quad (2)$$

The coefficient vector  $\beta_i(t)$  that resulted in the least-mean square difference between the levels of activity predicted by this model and the observed  $H\gamma(t)$  across all syllables was found by linear regression. For each electrode  $e_i$  at time  $t$ , the associated  $1 \times 4$  slope vector ( $\beta_i(t)$ ) quantifies the degree to which the engagement of a given articulator modulated the cross-syllabic variability in  $H\gamma(t)$  at that electrode. Coefficients of determination ( $R^2$ ) were calculated from the residuals of this regression. In the current context,  $R^2$  can be interpreted as the amount of cross-syllabic variability in  $H\gamma$  that can be explained by the optimally weighted linear combination of articulatory state variables.

The spatial organization of the speech articulators was examined using the assigned weight vectors ( $\beta_i(t)$ ) from the GLM described above. First, the fit of the GLM at each electrode  $e_i$  was determined to be of interest if, on average, the associated  $P$ -value was less than 0.05 for any one of the four consonant articulator time windows ( $T_A$ ; see below) determined from the partial-correlation analysis below. We defined this time window to be the average onset-to-offset time of statistically significant ( $P < 0.05$ ) partial correlations for each individual articulator in each subject (see the section on partial correlation analysis below). This method identifies electrodes whose activity is predicted well by the GLM for any of the individual articulators, as well as for combinations of these articulators, for extended periods of time. As these time windows extend for many points, this is a relatively stringent criterion in comparison to a min-finding method or looking for single significant-crossings. In practice, the minimum  $P$ -values (across time) associated with the vast majority of these electrodes are several orders of magnitude less than 0.05. For the electrodes that were gauged to have statistically significant correlations in each subject, we averaged the weights for each articulator ( $A$ ) in that articulators time window ( $T_A$ ). Thus, each electrode of interest ( $e_i$ ) is assigned four values, with each value corresponding to the weighting for that articulator, averaged across that articulator's time window:

$$W_i^A = \frac{1}{|T_A|} \sum_{t \in T_A} \beta_i(t) \quad (3)$$

For the analysis of representational overlap at individual electrodes (Supplementary Fig. 6), each electrode was classified according to the dominant articulator weight in a winner-take-all manner. The fractional articulator weighting was calculated based on the positive weights at each electrode, and was plotted as the average percentage of summed positive weights.

For spatial analysis, the data for each subject were smoothed using a 2-mm uniform circular kernel. The maps presented and analysed in Supplementary Figs 2 and 3 correspond to these average weights for the lips, tongue, larynx and jaw. The maps presented and analysed in Fig. 2 correspond to these average weights for each articulator averaged across subjects. The spatial organization of vSMC is described by plotting the results of the GLM for an individual on the cortex of that individual. We used a Cartesian plane defined by the antero-posterior distance from the central sulcus (ordinate) and the dorsoventral distance from the Sylvian fissure (azimuth). This provides a consistent reference frame to describe the spatial organization of each subject's cortex and to combine data across subjects while preserving the individual differences.

**Somatotopic map and  $k$ -nearest neighbours algorithm.** To construct the summary somatotopic map of Fig. 2b, we first extracted the spatial location of the top 10% of weights for each articulator (averaged across subjects, data are shown in Fig. 2a). We then used a  $k$ -nearest neighbour algorithm to classify the surrounding cortical tissue based on the nearest  $k = 4$  neighbours within a spatial extent of 3 mm of each spatial location; if no data points were present within 3 mm, the location is unclassified. Locations in which no clear majority ( $>50\%$ ) of the nearest neighbours belonged to a single articulator were classified as mixed. These values were chosen to convey, in summary form, the visual impression of the individual articulator maps, and to 'fill in' spatial gaps in our recordings.

The summary map changed smoothly, and as expected with changes in the threshold of weights for each articulator of individual articulator maps,  $k$  (number of neighbours), spatial extent and minimum number of points. Results are qualitatively insensitive to the details of this analysis, including the choice of 10% as a threshold, as changes in the clustering algorithm could be made to accommodate subtle differences in data inclusion (for visual comparison, we display the somatotopic maps derived from the same algorithm derived from the top 5%, top 10% and top 15% of weights in Supplementary Fig. 4).

**Partial correlation analysis.** To quantify the temporal structure with which single-electrode  $H\gamma$  was correlated with the engagement of a single articulator, we used partial correlation analysis. Partial correlation analysis is a standard statistical tool that quantifies the degree of association between two random variables (here,  $H\gamma(t)$  and the engagement of a given articulator,  $A_i$ ), and removes the effect of a set of other random variables (here, the other articulators,  $A_j, j \neq i$ ). For a given electrode, the partial correlation coefficient between  $H\gamma(t)$  across syllables at time  $t$  and articulator  $A_i$  ( $\rho(H\gamma(t), A_i)$ ) is calculated as the correlation coefficient between the residuals  $r(H\gamma(t), A_j), j \neq i$  resulting from de-correlating the  $H\gamma(t)$  and every other articulator  $A_j, j \neq i$ , and the residuals  $r(A_i, A_j), i \neq j$  resulting from de-correlating the articulators from one another:

$$\rho(H\gamma(t), A_i) = \frac{\text{cov}(r(H\gamma(t), A_j), r(A_i, A_j))}{\sigma_1 \times \sigma_2}, i \neq j \quad (4)$$

Where  $\sigma_1$  and  $\sigma_2$  are the standard deviations of  $r(H\gamma(t), A_j)$  and  $r(A_i, A_j)$ , respectively. In the current context, the partial correlation coefficients quantify the degree to which the cross-syllabic variability in  $H\gamma$  at a given moment in time was uniquely associated with the engagement of a given articulator during speech production. For each articulator, we analysed those electrodes whose peak partial correlation coefficient ( $\rho$ ) exceeded the mean  $\pm 2.5\sigma$  of  $\rho$  values across electrodes and time ( $> \text{mean}(\rho(e_i, t)) + 2.5\sigma(\rho(e_i, t))$ ). In the text, we focus on the positive correlations (which we denote as  $R$  in the text), because there were typically a larger number of positive values than negative values (mean  $\rho > 0$ ), and in general the temporal profiles are similar for negative values and for expositional simplicity. Results did not qualitatively change with changes in this threshold of approximately  $\pm 0.2\sigma$ . We extracted the onset, offset and peak times for each articulator for each electrode that crossed this threshold. The data presented in Fig. 3d are the mean  $\pm$  s.e.m. of these timing variables across electrodes pooled across subjects. The average onset and offset for each of the four consonant articulators (lips, tongue, jaw and larynx) in each subject was used to define the articulator time window used in the spatial analysis described above.

**Principal component analysis and cortical-state space.** Principal components analysis (PCA) was carried out on the set of all vSMC electrodes for dimensionality reduction and orthogonalization. PCA was performed on the  $n \times m \times t$  covariance matrix  $Z$  with rows corresponding to channels (of which there are  $n$ ) and columns corresponding to concatenated  $H\gamma(t)$  (length  $t$ ) for each consonant-vowel (of which there are  $m$ ). Each electrode's time series was z-scored across syllables to normalize response variability across electrodes. The singular-value decomposition of  $Z$  was used to find the eigenvector matrix  $M$  and associated eigenvalues  $\lambda$ . The principal components (PCs) derived in this way serve as a spatial filter of the electrodes, with each electrode  $e_j$  receiving a weighting in  $PC_i$  equal to  $M_{ij}$ , where  $M$  is the matrix of eigenvectors. Across subjects, we observed that the eigenvalues ( $\lambda$ ) exhibited a fast decay with a sharp inflection point at the ninth eigenvalue, followed by a much slower decay thereafter (Supplementary Fig. 8). We therefore used the first nine eigenvectors (PCs) as the cortical state-space for each subject.

The cortical state-space representation of syllable  $s_k$  at time  $t$ ,  $K(s_k, t)$ , is defined as the projection of the vector of cortical activity associated with  $s_k$  at time  $t$ ,  $H\gamma_k(t)$ , onto  $M$ :

$$K(s_k, t) = M \bullet H\gamma_k(t) \quad (5)$$

We calculated the contribution of articulators to the cortical state-space ( $PCw_{ij}$ ) by projecting each electrode's weight vector ( $\beta_j$ ; derived from the GLM model above) into the first three dimensions of the cortical state-space ( $i = 1-3$ ):

$$PCw_{ij} = M_{ij} \bullet \beta_j \quad (6)$$

Here,  $PCw_{ij}$  is a four-element vector of the projected articulator weights for electrode  $e_j$  into  $PC_i$ . In Supplementary Fig. 9, we plot  $\log_{10}$  of the absolute value of  $PCw_{ij}$  across electrodes, which describes the distribution of magnitudes of the representations associated with the four articulators in a given PC.

**Clustering analysis.** The  $k$ -means and hierarchical clustering analyses were carried out on the cortical state-space representations of syllables,  $K(s_k, t)$ , based on the pair-wise Euclidean distances calculated between consonant-vowel syllable representations. Agglomerative hierarchical clustering used Ward's method. All analyses of the detailed binary phonetic feature matrix were carried out using both Hamming and Euclidean distances; results did not change between metrics qualitatively or statistically. We used silhouette analysis to validate the claim that there were three clusters at the consonant time. The silhouette of a cluster is a measure of how close (on average) the members of that cluster are to each other, relative to the next nearest cluster. For a particular data set, the average silhouette for a given number of clusters describes the parsimony of the number of clusters in the data. Hence, examining the silhouette across different numbers of clusters gives a quantitative way to determine the most parsimonious number of clusters<sup>51</sup>. Higher values correspond to more parsimonious clustering. On average across subjects, this analysis validated the claim that three clusters (average silhouette = 0.47) was a more parsimonious clustering scheme than either two (average silhouette = 0.45) or four clusters (average silhouette = 0.43). **Correlation of cortical state-space structure with phonetic structure.** At each moment in time, we wanted to quantify the similarity of the structure of cortical state-space representations of phonemes and the structure predicted by different phonetic feature sets. To do this, we measured the linear correlation coefficient between vectors of unique pair-wise Euclidean distances between phonemes calculated in the cortical state-space ( $DC(t)$ ) and in the phonetic feature matrix ( $DP$ ):

$$R(t) = \frac{\text{cov}(DC(t), DP)}{\sigma_{DC(t)} \times \sigma_{DP}} \quad (7)$$

As described above, the phonetic feature matrix was composed of three distinct phonetic feature sets (consonant constriction location, consonant constriction degree or shape, vowel configuration). Distances were calculated independently in these three subsets and correlated with  $DC(t)$ . Standard error measures of the correlation coefficients were calculated using a bootstrap procedure (1,000 iterations).

**Cluster separability.** Cluster separability is defined at any moment in time as the difference between the average of cross-cluster distances and the average of within-cluster distances. This quantifies the average difference of the distance between syllables in different clusters and the tightness of a given cluster. We quantified the variability in cluster separability estimation using a 1,000-iteration bootstrap procedure of the syllables used to calculate the metric.

**Cluster density.** We quantified the average cluster density by calculating the average inverse of all unique pair-wise distances between consonant-vowels in a given cortical state-space cluster. It is a density because the number of elements in a cluster does not change with time.

48. Mesgarani, N. & Chang, E. F. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* **485**, 233–236 (2012).
49. Yang, X. et al. Auditory representations of acoustic signals. *IEEE Transactions Inf. Theor.* **38**, 824–839 (1992).
50. International Phonetic Association. *Handbook of the International Phonetic Association* (Cambridge Univ. Press, 1999).
51. Rousseeuw, P. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).



# Circular RNAs are a large class of animal RNAs with regulatory potency

Sebastian Memczak<sup>1\*</sup>, Marvin Jens<sup>1\*</sup>, Antigoni Elefanti<sup>1\*</sup>, Francesca Torti<sup>1\*</sup>, Janna Krueger<sup>2</sup>, Agnieszka Rybak<sup>1</sup>, Luisa Maier<sup>1</sup>, Sebastian D. Mackowiak<sup>1</sup>, Lea H. Gregersen<sup>3</sup>, Mathias Munschauer<sup>3</sup>, Alexander Loewer<sup>4</sup>, Ulrike Ziebold<sup>1</sup>, Markus Landthaler<sup>3</sup>, Christine Kocks<sup>1</sup>, Ferdinand le Noble<sup>2</sup> & Nikolaus Rajewsky<sup>1</sup>

Circular RNAs (circRNAs) in animals are an enigmatic class of RNA with unknown function. To explore circRNAs systematically, we sequenced and computationally analysed human, mouse and nematode RNA. We detected thousands of well-expressed, stable circRNAs, often showing tissue/developmental-stage-specific expression. Sequence analysis indicated important regulatory functions for circRNAs. We found that a human circRNA, antisense to the cerebellar degeneration-related protein 1 transcript (CDR1as), is densely bound by microRNA (miRNA) effector complexes and harbours 63 conserved binding sites for the ancient miRNA miR-7. Further analyses indicated that CDR1as functions to bind miR-7 in neuronal tissues. Human CDR1as expression in zebrafish impaired midbrain development, similar to knocking down miR-7, suggesting that CDR1as is a miRNA antagonist with a miRNA-binding capacity ten times higher than any other known transcript. Together, our data provide evidence that circRNAs form a large class of post-transcriptional regulators. Numerous circRNAs form by head-to-tail splicing of exons, suggesting previously unrecognized regulatory potential of coding sequences.

Mature messenger RNAs are linear molecules with 5' and 3' termini that reflect start and stop of the RNA polymerase on the DNA template. In cells, different RNA molecules are sometimes joined together by splicing reactions (trans-splicing), but covalent linkage of the ends of a single RNA molecule to form a circular RNA (circRNA) is usually considered a rare event. circRNAs were discovered in plants and shown to encode subviral agents<sup>1</sup>. In unicellular organisms, circRNAs mostly stem from self-splicing introns of pre-ribosomal RNA<sup>2</sup>, but can also arise from protein-coding genes in archaea<sup>3</sup>. In the few unambiguously validated circRNAs in animals, the spliceosome seems to link the 5' and downstream 3' ends of exons within the same transcript<sup>4–10</sup>. Perhaps the best known circRNA is antisense to the mRNA transcribed from the SRY (sex-determining region Y) locus and is highly expressed in testes<sup>6</sup>. Evidence from computational analyses of expression data in Archaea and Mammalia suggests that circRNAs are more prevalent than previously thought<sup>3,10</sup>; however, it is unknown whether animal circRNAs have any biological function.

In comparison to circRNAs, miRNAs are extremely well studied. miRNAs are ~21-nucleotide-long non-coding RNAs that guide the effector protein Argonaute (AGO) to mRNAs of coding genes to repress their protein production<sup>11–14</sup>. In humans, miRNAs directly regulate expression of most mRNAs<sup>15–18</sup> in a diverse range of biological functions. However, surprisingly little is known about how and if mRNAs can escape regulation by a miRNA. A recently discovered mechanism for miRNA removal in a sequence-specific manner is based on target sites acting as decoys or miRNA sponges<sup>19,20</sup>. RNA with miRNA binding sites should, if expressed highly enough, sequester away the miRNA from its target sites. However, all reported mammalian miRNA sponges have only one or two binding sites for the same miRNA and are not highly expressed, limiting their potency<sup>21–24</sup>.

To identify circRNAs across animal cells systematically, we screened RNA-seq data for circRNAs. Compared to previous approaches<sup>10</sup> our

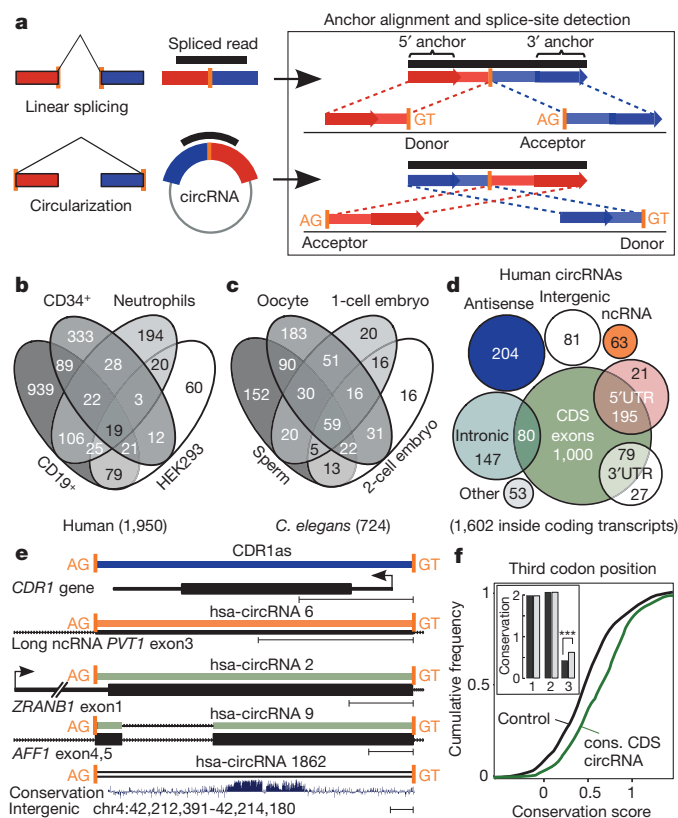
computational pipeline can find circRNAs in any genomic region, takes advantage of long (~100 nucleotides) reads, and predicts the acceptor and donor splice sites used to link the ends of the RNAs. We do not rely on paired-end sequencing data or known splice sites. Using published<sup>10,25,26</sup> and our own sequencing data, our method reported thousands of circRNAs in human and mouse tissues as well as in different developmental stages of *Caenorhabditis elegans*. Numerous circRNAs appear to be specifically expressed across tissues or developmental stages. We validated these data and showed that most tested circRNAs are well expressed, stable and circularized using the predicted splice sites. circRNA sequences were significantly enriched in conserved nucleotides, indicating that circRNAs compete with other RNAs for binding by RNA binding proteins (RBPs) or miRNAs. We combined biochemical, functional and computational analyses to show that indeed a known human circRNA, CDR1 antisense (CDR1as)<sup>9</sup>, can function as a negative regulator of miR-7, a miRNA with perfect sequence conservation from annelids to human. Together, our data provide evidence that circRNAs form an important class of post-transcriptional regulators.

## circRNAs have complex expression patterns

To comprehensively identify stably expressed circRNAs in animals we screened RNA sequencing reads for splice junctions formed by an acceptor splice site at the 5' end of an exon and a donor site at a downstream 3' end (head-to-tail) (Fig. 1a). As standard RNA expression profiling enriches for polyadenylated RNAs, we used data generated after ribosomal RNA depletion (ribominus) and random priming. Such data were used before to detect scrambled exons in mammals<sup>10</sup> (see Methods for comparison). However, this approach was not specifically designed to detect circRNAs and (1) only used existing exon–intron annotations, thus missing RNAs transcribed from introns or unannotated transcripts; (2) did not explicitly identify

<sup>1</sup>Systems Biology of Gene Regulatory Elements, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany. <sup>2</sup>Angiogenesis and Cardiovascular Pathology, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany. <sup>3</sup>RNA Biology and Post-Transcriptional Regulation, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany. <sup>4</sup>Signaling Dynamics in Single Cells, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany.

\*These authors contributed equally to this work.



**Figure 1 | Detection, classification and evolutionary conservation of circRNAs.** **a**, The termini of junction-spanning reads (anchors) align sequentially to the genome for linear (top) but in reversed orientation for head-to-tail spliced reads (bottom). Spliced reads must distribute completely to anchors, flanked by AG/GU (Methods). **b**, **c**, circRNAs in human cell types (**b**) and nematode stages (**c**). **d**, Genomic origin of human circRNAs. A total of 96% of circRNAs overlap known transcripts. **e**, Examples of human circRNAs. The *AFF1* intron is spliced out (Supplementary Fig. 2e). Sequence conservation: placental mammals phyloP score (Methods), scale bar, 200 nucleotides. **f**, A total of 223 human coding sequence circRNAs with mouse orthologues (green) and controls (black) with matched conservation level (inset: mean conservation for each codon position (grey), controls (black); x axis, codon positions; y axis, placental mammals phyloP score; see also Methods and Supplementary Fig. 1j, k). Third codon positions are significantly more conserved ( $P < 4 \times 10^{-10}$ , Mann–Whitney *U*-test,  $n = 223$ ).

the splice sites used for circularization; and (3) assumed that each pair of mates in paired-end sequencing derives from the same RNA molecule. To search in a more unbiased way for circRNAs, we designed an algorithm (Methods) that identifies linear and circular splicing events in ribominus data. First, we filtered out reads that aligned contiguously to the genome, retaining the spliced reads. Next, we mapped the terminal parts of each candidate read independently to the genome to find unique anchor positions. Finally, we demanded that (1) anchor alignments can be extended such that the original read sequence aligns completely, and (2) the inferred breakpoint is flanked by GU/AG splice signals. Non-unique mappings and ambiguous breakpoints were discarded. We detected circularization splicing from the reversed (head-to-tail) orientation of the anchor alignments (Fig. 1a). Our method also recovered tens of thousands of known linear splicing events (Methods and Supplementary Fig. 1a, b). We estimated sensitivity ( $>75\%$ ) and false-discovery rate (FDR  $<0.2\%$ ) using simulated reads and various permutations of real sequencing data (Methods and Supplementary Fig. 1c). However, the efficiency of ribominus protocols to extract and sequence circRNAs is limited, reducing overall sensitivity.

We generated ribominus data for HEK293 cells and, combined with human leukocyte data<sup>10</sup>, detected 1,950 circRNAs with support

from at least two independent junction-spanning reads (Fig. 1b). The expression of genes predicted to give rise to circRNAs was only slightly shifted towards higher expression values (Supplementary Fig. 1d), indicating that circRNAs are not just rare mistakes of the spliceosome. We also identified 1,903 circRNAs in mouse (brains, fetal head, differentiation-induced embryonic stem cells; Supplementary Fig. 1e)<sup>25,26</sup>; 81 of these mapped to human circRNAs (Supplementary Fig. 1f). To explore whether circRNAs exist in other animal clades, we used sequencing data that we produced from various *C. elegans* developmental stages (Stoeckius, M. *et al.*, manuscript in preparation) (Methods) and detected 724 circRNAs, with at least two independent reads (Fig. 1c).

Numerous circRNAs seem to be specifically expressed in a cell type or developmental stage (Fig. 1b, c and Supplementary Fig. 1e). For example, hsa-circRNA 2149 is supported by 13 unique, head-to-tail spanning reads in CD19<sup>+</sup> leukocytes but is not detected in CD34<sup>+</sup> leukocytes (which were sequenced at comparable depth; Supplementary Table 1), neutrophils or HEK293 cells. Analogously, a number of nematode circRNAs seem to be expressed in oocytes but absent in 1- or 2-cell embryos.

We annotated human circRNAs using the RefSeq database and a catalogue of non-coding RNAs<sup>27–29</sup>. 85% of human circRNAs align sense to known genes. Their splice sites typically span one to five exons (Supplementary Fig. 1g) and overlap coding exons (84%), but only in 65% of these cases are both splice sites that participate in the circularization known splice sites (Supplementary Table 2), demonstrating the advantage of our strategy. 10% of all circRNAs align antisense to known transcripts, smaller fractions align to UTRs, introns, unannotated regions of the genome (Fig. 1d). Examples of human circRNAs are shown in Fig. 1e.

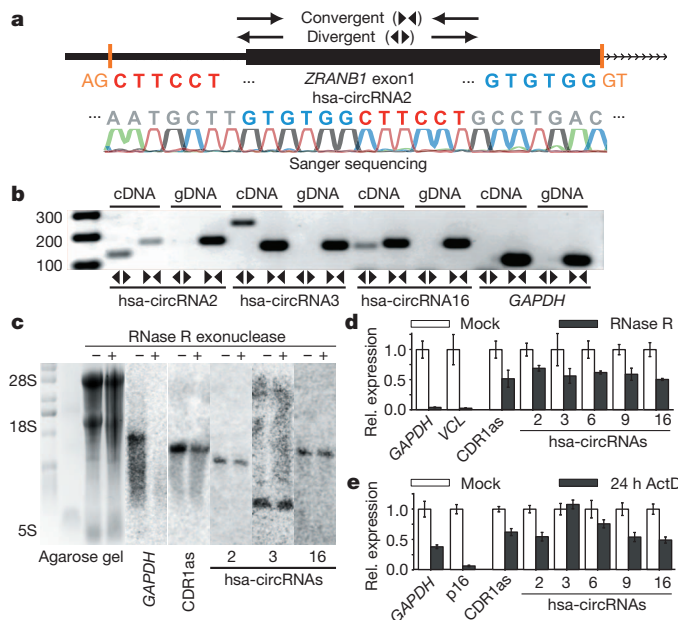
We analysed sequence conservation within circRNAs. As genomic sequence is subject to different degrees of evolutionary selection, depending on function, we studied three subtypes of circRNAs. Intergenic and a few intronic circRNAs display a mild but significant enrichment of conserved nucleotides (Supplementary Fig. 1h, i). To analyse circRNAs composed of coding sequence and thus high overall conservation, we selected 223 human circRNAs with circular orthologues in mouse (Methods) and entirely composed of coding sequence. Control (linear) exons were randomly selected to match the level of conservation observed in first and second codon positions (Methods, Fig. 1f inset and Supplementary Fig. 1k for conservation of the remaining coding sequence (CDS)). circRNAs with conserved circularization were significantly more conserved in the third codon position than controls, indicating evolutionary constraints at the nucleotide level, in addition to selection at the protein level (Fig. 1f and Supplementary Fig. 1j, k). In summary, we have confidently identified a large number of circRNAs with complex expression patterns, which derive often but not always from coding exons. Sequence conservation suggests that at least a subset contains functional sequence elements.

## Characterization of 50 predicted circRNAs

We experimentally tested our circRNA predictions in HEK293 cells. Head-to-tail splicing was assayed by quantitative polymerase chain reaction (qPCR) after reverse transcription, with divergent primers and Sanger sequencing (Fig. 2a, b). Predicted head-to-tail junctions of 19 out of 23 randomly chosen circRNAs (83%) could be validated, demonstrating high accuracy of our predictions (Table 1). In contrast, 5 out of 7 (71%) candidates exclusively predicted in leukocytes could not be detected in HEK293 cells, validating cell-type-specific expression.

Head-to-tail splicing could be produced by trans-splicing or genomic rearrangements. To rule out these possibilities as well as potential PCR artefacts, we successfully validated the insensitivity of human circRNA candidates to digestion with RNase R—an exonuclease that degrades linear RNA molecules<sup>30</sup>—by northern blotting with probes which span the head-to-tail junctions (Fig. 2c). We quantified RNase R resistance for 21 candidates with confirmed head-to-tail splicing by





**Figure 2 | CircRNAs are stable transcripts with robust expression.** **a**, Human (hsa) *ZRANB1* circRNA exemplifies the validation strategy. Convergent (divergent) primers detect total (circular) RNAs. Sanger sequencing confirms head-to-tail splicing. **b**, Divergent primers amplify circRNAs in cDNA but not genomic DNA (gDNA). *GAPDH*, linear control, size marker in base pairs. **c**, Northern blots of mock (–) and RNase R (+) treated HEK293 total RNA with head-to-tail specific probes for circRNAs. *GAPDH*, linear control. **d**, **e**, circRNAs are at least 10-fold more RNase R resistant than *GAPDH* mRNA (**d**) and stable after 24 h transcription block (**e**) (qPCR; error bars indicate standard deviation).

qPCR. All of these were at least 10-fold more resistant than *GAPDH* (Fig. 2d and Supplementary Fig. 2a). We reasoned that circRNAs should generally turn over more slowly than mRNAs. Indeed, we found that 24 h after blocking transcription circRNAs were highly stable, exceeding the stability of the housekeeping gene *GAPDH*<sup>31</sup> (Fig. 2e and Supplementary Fig. 2b). We also validated 3 out of 3 tested mouse circRNAs with human orthologues in mouse brains (Supplementary Fig. 2c). In *C. elegans* 15 out of 20 (75%) of the predictions from gametes and early embryos were validated in a mixed stage sample (Supplementary Fig. 2d and Supplementary Table 3).

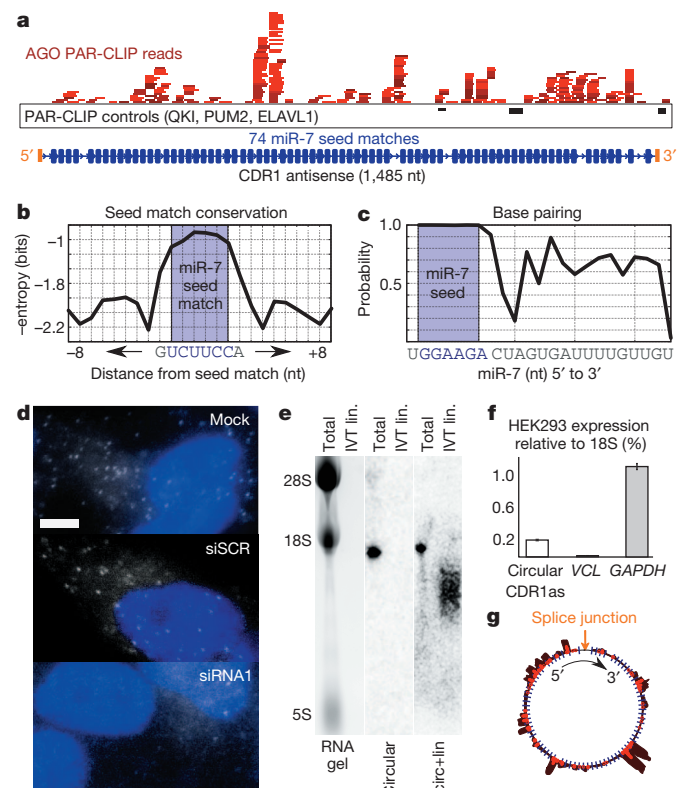
### circRNA CDR1as is densely bound by AGO

Stable transcripts with many miRNA-binding sites could function as miRNA sponges. We intersected our catalogue of circRNAs with transcript annotations, assuming that introns would not occur in mature circRNAs (as observed for 3 out of 3 tested circRNAs, Supplementary Fig. 2e). We screened for occurrences of conserved miRNA family seed matches (Methods). When counting repetitions of conserved matches to the same miRNA family, circRNAs were

significantly enriched compared to coding sequences ( $P < 2.96 \times 10^{-22}$ , Mann-Whitney *U*-test,  $n = 3,873$ ) or 3' UTR sequences ( $P < 2.76 \times 10^{-21}$ , Mann-Whitney *U*-test,  $n = 3,182$ ) (Supplementary Fig. 3a, b).

As an extreme case, we discovered that the known human circRNA CDR1as (ref. 9) harboured dozens of conserved miR-7 seed matches. To test whether CDR1as is bound by miRNAs, we analysed biochemical, transcriptome-wide binding-site data for the miRNA effector AGO proteins. We performed four independent PAR-CLIP (photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation) experiments for human AGO (Methods) and analysed them together with published, lower-depth data<sup>32</sup>. PAR-CLIP<sup>32–34</sup> is based on ultraviolet crosslinking of RNA to protein and subsequent sequencing of RNA bound to a RBP of interest. The ~1.5-kilobase (kb) *CDR1as* locus stood out in density and number of AGO PAR-CLIP reads (Fig. 3a), whereas nine combined PAR-CLIP libraries for other RBPs gave virtually no signal. Of note, there is no PAR-CLIP read mapping to the sense coding transcript of the *CDR1* gene, which was originally identified as a target of autoantibodies from patients with paraneoplastic cerebellar degeneration<sup>35</sup>.

Sequence analysis across 32 vertebrate species revealed that miR-7 is the only animal miRNA with conserved seed matches that can explain the AGO binding along the CDR1as transcript (Methods). Human CDR1as harbours 74 miR-7 seed matches of which 63 are



**Figure 3 | The circRNA CDR1as is bound by the miRNA effector protein AGO, and is cytoplasmic.** **a**, CDR1as is densely bound by AGO (red) but not by unrelated proteins (black). Blue boxes indicate miR-7 seed matches. nt, nucleotides. **b**, **c**, miR-7 sites display reduced nucleotide variability across 32 vertebrate genomes (**b**) and high base-pairing probability within seed matches (**c**). **d**, CDR1as RNA is cytoplasmic and disperse (white spots; single-molecule RNA FISH; maximum intensity merges of Z-stacks). siSCR, positive; siRNA1, negative control. Blue, nuclei (DAPI); scale bar, 5  $\mu$ m (see also Supplementary Fig. 10 for uncropped images). **e**, Northern blotting detects circular but not linear CDR1as in HEK293 RNA. Total, HEK293 RNA; circular, head-to-tail probe; circ+lin, probe within splice sites; IVT lin., *in vitro* transcribed, linear CDR1as RNA. **f**, Circular CDR1as is highly expressed (qPCR, error bars indicate standard deviation). **g**, CDR1as. Blue, seed matches; dark red, AGO PAR-CLIP reads; bright red, crosslinked nucleotide conversions.

**Table 1 | Summary of the validation experiments**

Sample	Validation experiment	Validation success
Human (HEK293)	Head-to-tail splicing	19 of 23
	Circularity	21 of 21
	Expression >3% vinculin	12 of 21
	Expression specificity (leukocyte specific)	5 of 7
Mouse (adult brain)	Head-to-tail splicing	3 of 3
	Circularity	3 of 3
	Expression >1% $\beta$ -actin	2 of 3
<i>C. elegans</i>	Head-to-tail splicing	15 of 20
	Circularity	13 of 13
	Expression >1% <i>elf-3.d</i>	12 of 15

Most experimentally tested circRNAs are validated.

conserved in at least one other species (Supplementary Fig. 4). Interspaced sequences were less conserved, indicating that miR-7 binding sites are probably functional (Fig. 3b). Secondary structure analysis of predicted circRNA-miRNA duplexes (Methods) showed reduced base-pairing of miR-7 beyond the seed (Fig. 3c). None of the ~1,500 miR-7 complementary sites across 32 vertebrate sequences was complementary beyond position 12 of miR-7 (only three could form an 11-nucleotide duplex) (Supplementary Table 4). Slicing by mammalian Argonaute requires complementarity of positions 10 and 11 and depends on extended complementarity beyond position 12 (ref. 36). Thus, CDR1as seems optimized to be densely bound but not sliced by miR-7.

Single-molecule imaging (Methods) revealed disperse and mostly cytoplasmic CDR1as expression (HEK293 cells), consistent with miRNA sponge function (Fig. 3d and Supplementary Table 5). CDR1as circularization was assayed by northern blotting (Fig. 3e). Nicking experiments confirmed that CDR1as circRNA can be linearized and degraded (Supplementary Fig. 5a). In RNA from HEK293 cells, circularized but no additional linear CDR1as was detected (Supplementary Fig. 5b). Circular expression levels were quantified by qPCR with divergent primers calibrated by standard curves (Supplementary Table 6). CDR1as was highly expressed (~15% to ~20% of *GAPDH* expression, Fig. 3f). Estimating *GAPDH* mRNA copy number from HEK293 RNA-seq data (~1,400 molecules per cell, data not shown) suggests that CDR1as may bind up to ~20,000 miR-7 molecules per cell (Fig. 3g).

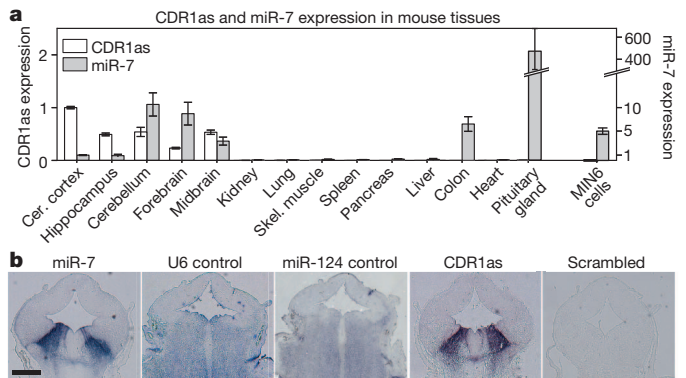
If CDR1as functions as a miR-7 sponge, its destruction could trigger downregulation of miR-7 targets. We knocked down CDR1as in HEK293 cells and monitored expression of published miR-7 targets by qPCR with externally spiked-in standards (Methods and Supplementary Fig. 5c, d). All eight miR-7 targets assayed, but also housekeeping genes, were downregulated. Nanostring technology<sup>37</sup> additionally indicated downregulation of many genes (data not shown). Furthermore, stable loss of CDR1as expression by virally delivered small hairpin RNAs led to significantly reduced migration in an *in vitro* wound closure assay (Methods, Supplementary Fig. 5e, f and Supplementary Table 7). Thus, knockdown of CDR1as affects HEK293 cells, but we could not delineate miR-7-specific effects, potentially because of indirect or miR-7-independent CDR1as function (see below).

### Co-expression of miR-7 and CDR1as in brain

If CDR1as indeed interacts with miR-7, both must be co-expressed. miR-7 is highly expressed in neuronal tissues, pancreas and pituitary gland<sup>38</sup>. Apart from HEK293 cells, a cell line probably derived from neuronal precursors in embryonic kidney<sup>39</sup>, we quantified miR-7 and CDR1as expression across mouse tissues and pancreatic-island-derived MIN6 cells (Methods and Fig. 4a). CDR1as and miR-7 were both highly expressed in brain tissues, but CDR1as was expressed at low levels or absent in non-neuronal tissues, including tissues with very high miR-7 expression. qPCR suggested that CDR1as is exclusively circular in adult and embryonic mouse brain (Supplementary Fig. 5g, h). Thus, CDR1as and miR-7 seem to interact specifically in neuronal tissues. Indeed, when assaying CDR1as and miR-7 in mouse brains by *in situ* hybridizations (Methods), we observed specific, similar, but not identical, expression patterns in the brain of mid-gestation (embryonic day 13.5 (E13.5)) embryos (Fig. 4b). Specifically, CDR1as and miR-7 were highly co-expressed in areas of the developing midbrain (mesencephalon)<sup>40,41</sup>. Thus, CDR1as is highly expressed, stable, cytoplasmic, not detectable as a linear RNA and shares expression domains with miR-7. Together with extensive miR-7 binding within CDR1as, CDR1as has hallmarks of a potent circular miR-7 sponge in neuronal tissues.

### Effects of miR-7 and CDR1as in zebrafish

It would be informative to knock out CDR1as in an animal model system. However, a knockout would also affect CDR1 protein, with



**Figure 4 | CDR1as and miR-7 have overlapping and specific expression in neuronal tissues.** **a**, Among mouse tissues and MIN6 cells (qPCR, relative to cerebral cortex expression; error bars indicate standard deviations; see Supplementary Fig. 9a for miR-122 control) neuronal tissues co-express miR-7 and CDR1as. **b**, *In situ* staining of CDR1as and miR-7 in mouse embryo brain E13.5 (U6 and miR-124, positive control; scrambled probe, negative control). Scale bar, 1 mm.

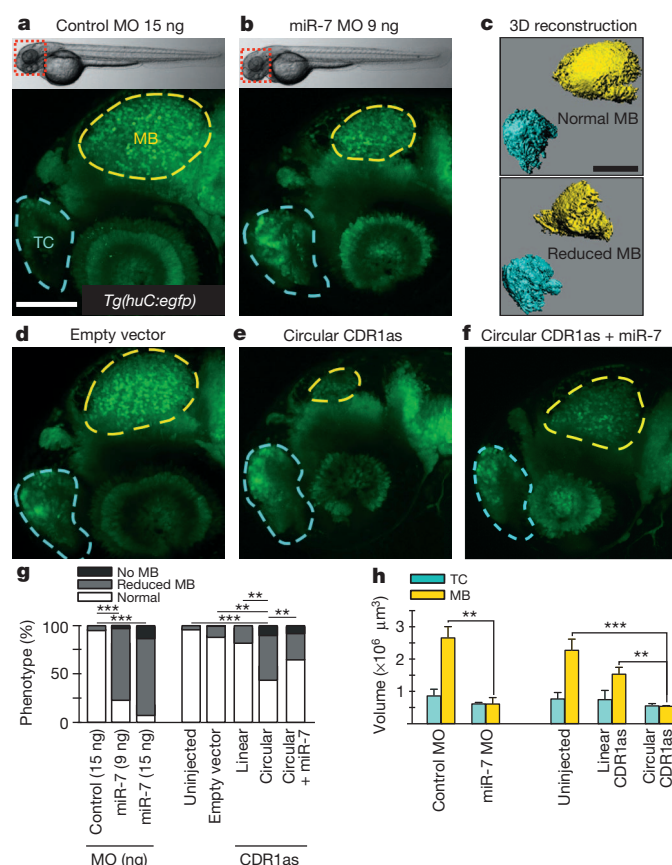
unknown consequences. This problem is circumvented when using zebrafish (*Danio rerio*) as an animal model. According to our bioinformatic analyses (not shown) zebrafish has lost the *cdr1* locus, whereas miR-7 is conserved and highly expressed in the embryonic brain<sup>42</sup>. Thus, we can test whether miR-7 has a loss-of-function phenotype and if this phenotype can be induced by introduction of mammalian CDR1as RNA. We injected morpholinos to knock down mature miR-7 expression in zebrafish embryos (Methods). At a dose of 9 ng of miR-7 morpholino, the embryos did not show overall morphological defects but reproducibly, and in two independent genetic backgrounds (Supplementary Fig. 6a–c), developed brain defects (Fig. 5a, b). In particular, ~70% showed a consistent and clear reduction in midbrain size, and an additional ~5% of animals had almost completely lost their midbrains. Of note, the telencephalon at the anterior tip of the brain was not affected in size. Brain volumes were also measured based on confocal three-dimensional stacks (Fig. 5c and Supplementary Fig. 7). Reduction of the midbrain size correlated with miR-7 inhibition in the respective animals (Supplementary Fig. 6d). These data provide evidence that miR-7 loss-of-function causes a specific reduction of midbrain size.

To test whether CDR1as can function as a miR-7 sponge *in vivo*, we injected embryos with plasmid DNA that expressed a linear version of the full-length human *CDR1as* sequence (Supplementary Fig. 6e, f) or a plasmid provided by the Kjemis laboratory that can produce circular CDR1as in human cells (Fig. 5d, e). qPCR analysis detected circular RNA in zebrafish embryos injected with the latter plasmid (Supplementary Fig. 8), which reproducibly and in independent genetic backgrounds lead to reduced midbrain sizes (Fig. 5g, h). Similarly, animals injected with *in vitro*-transcribed partial mouse CDR1as RNA, but not with RNA from the other strand, showed significant midbrain reduction (Supplementary Fig. 6g–i). Thus, the phenotype is probably caused by CDR1as RNA and not by an unspecific effect of RNA or DNA injection. These results provide evidence that human/mouse CDR1as transcripts are biologically active *in vivo* and impair brain development similarly to miR-7 inhibition. The midbrain reduction could be partially rescued by injecting miR-7 precursor (Fig. 5f, g), arguing that the biological effect of CDR1as expression is caused at least in part by interaction of CDR1as with miR-7.

### Discussion

We have shown that animal genomes express thousands of circRNAs from diverse genomic locations (for example, from coding and non-coding exons, intergenic regions or transcripts antisense to 5' and 3' UTRs) in a complex tissue-, cell-type- or developmental-stage-specific manner. We provided evidence that CDR1as can act as a





**Figure 5 | In zebrafish, knockdown of miR-7 or expression of CDR1as causes midbrain defects.** **a, b**, Neuronal reporter (*Tg(huC:egfp)*) embryos (top, light microscopy) 48 h post fertilization (bottom, representative confocal z-stack projections; blue dashed line, telencephalon (TC) (control); yellow dashed line, midbrain (MB)). Embryos after injection of 9 ng miR-7 morpholino (MO) (**b**) display a reduction in midbrain size. Panel **a** shows a representative embryo injected with 15 ng control morpholino. **c**, Three-dimensional volumetric reconstructions. **d**, Empty vector control. **e**, Expression vector encoding human circular CDR1as. **f**, Rescue experiment with miR-7 precursor. **g**, Phenotype penetrance (% of embryos, miR-7 MO,  $n = 135$ ; uninjected,  $n = 83$ ; empty vector,  $n = 91$ ; linear CDR1as,  $n = 258$ ; circular CDR1as,  $n = 153$ ; circular CDR1as plus miR-7 precursor,  $n = 217$ ). Phenotype distribution derived from at least three independent experiments. Scale bar, 0.1 mm.  $^{**}P < 0.01$ ;  $^{***}P < 0.001$  in Student's *t*-test for normal midbrain, reduced midbrain (see also Supplementary Fig. 6). **h**, Phenotype quantification (Methods). Error bars indicate standard deviation  $n = 3$  per group.

post-transcriptional regulator by binding miR-7 in brain tissues: (1) CDR1as is densely bound by miRNA effector molecules; (2) CDR1as harbours 74 miR-7 seed matches, often deeply conserved; (3) CDR1as is expressed highly, stably and mostly cytoplasmic; (4) CDR1as and miR-7 share specific expression domains in mouse embryonic brain; (5) human/mouse CDR1as is circularized *in vivo* and is not detectable as a linear molecule; (6) human/mouse CDR1as sequences, when injected into zebrafish, and miR-7 knock down have similar phenotypes in brain. While zebrafish circularization of human CDR1as may be incomplete, the midbrain phenotype was stronger compared to expressing linear CDR1as RNA that lacks circularization splice sites. Although the two DNA plasmids used carry identical promoters and were injected in equal concentrations, we cannot rule out the possibility that the difference in midbrain phenotype strength may be explained by other factors. However, because of the observed extreme stability of CDR1as and circRNAs in general, our data argue that circRNAs can be used as potent inhibitors of miRNAs or RBPs. Future studies should elucidate how CDR1as can be converted into a linear molecule and targeted for degradation. miR-671 can trigger destruction of

CDR1as<sup>9</sup>. Thus, CDR1as may function to transport miR-7 to subcellular locations, where miR-671 could trigger release of its cargo. Known functions of miR-7 targets such as PAK1 and FAK1 support these speculations<sup>43,44</sup>.

The phenotype induced by CDR1as expression in zebrafish was only partially rescued by expressing miR-7, indicating that CDR1as could have functions beyond sequestering miR-7. This idea is supported by *in situ* hybridization in mouse adult hippocampus (Supplementary Fig. 9b) where areas staining for CDR1as but not miR-7 were observed. What could be additional functions of circRNAs beyond acting as sponges? As a single-stranded RNA, CDR1as could, for example, bind *in trans* 3' UTRs of target mRNAs to regulate their expression. It is even possible that miR-7 binds CDR1as to silence these trans-acting activities. Alternatively, CDR1as could be involved in the assembly of larger complexes of RNA or protein, perhaps similar to other low-complexity molecules<sup>45</sup>.

How many other circRNAs exist? In this study, we identified approximately 2,000 human, 1,900 mouse and 700 nematode circRNAs from sequencing data, and our validation experiments confirmed most of the 50 tested circRNAs. However, we analysed only a few tissues/developmental stages with stringent cutoffs. Thus, the true number of circRNAs is almost certainly much larger. Although CDR1as is an extreme case, many circRNAs have conserved seed matches. For example, circRNA from the SRY locus<sup>6</sup> has seed sites for murine miRNAs. Therefore, circRNAs probably compete with other RNAs for miRNA binding. Sequence analyses indicated that coding exons serve additional, presumably regulatory functions when expressed within circRNAs, whereas intergenic or intronic circRNAs generally showed only weak conservation. Because we detected thousands of circRNAs, it is appealing to speculate that occasional circularization of exons is easy to evolve and may provide a mechanism for rapid evolution of stably and well expressed regulatory RNAs. Of note, we detected multiple seed matches for viral miRNAs within human circRNAs (not shown). However, there is no reason to think that circRNAs function predominantly to bind miRNAs. As known in bacteria, the decoy mechanism underlying miRNA sponges could be important also for RBPs<sup>46,47</sup>. Similarly, circRNAs could function to store, sort, or localize RBPs. In summary, our data suggest that circRNAs form a class of post-transcriptional regulators which compete with other RNAs for binding by miRNAs and RBPs and may generally function in modulating the local free concentration of RBPs, RNAs, or their binding sites.

**Note added in proof:** While this paper was under review, circular RNAs in fibroblasts were described<sup>51</sup>.

## METHODS SUMMARY

**Computational pipeline for predicting circRNAs from ribominus sequencing data.** A detailed description of the computational methods is given in the Methods.

**Cell culture and treatments.** HEK293, HEK293TN and HEK293 Flp-In 293 T-REx (Life Technologies) were cultured following standard protocols. Transcription was blocked by adding 2 μg ml<sup>-1</sup> actinomycin D (Sigma). RNase R (Epicentre Biotechnologies) treatment (3 U μg<sup>-1</sup>) was performed on total RNA (5 μg) at 37 °C for 15 min. qPCR primers are listed in Supplementary Table 8.

**Single-molecule RNA fluorescence *in situ* hybridization (smRNA FISH).** Stellaris Oligonucleotide probes complementary to CDR1as were designed using the Stellaris Probe Designer (Biosearch Technologies). Probe pools were obtained from BioCat GmbH as conjugates coupled to Quasar 670. Probes were hybridized at 125 nM at 37 °C. Images were acquired on an inverted Nikon Ti microscope.

**Mouse strains and *in situ* hybridization.** *In situ* hybridization (ISH) was performed on paraffin tissue sections from B6129SF1/J wild-type mice as described<sup>48</sup> using locked nucleic acid (LNA) probes or RNAs obtained by *in vitro* transcription on PCR products.

**Zebrafish methods.** *Tg(huC:egfp)* and *Tg(Xia.Tubb:dsRED)* transgenic zebrafish lines were used<sup>49,50</sup>. Morpholino antisense oligomers were injected into the yolk of single-cell-stage embryos. Furthermore, two pCS2+ plasmids coding for full-length linear CDR1as or CDR1as plus upstream and downstream sequence that can express circular CDR1as in human cells (courtesy of the Kjems laboratory)

were injected. Confocal imaging was performed using Carl Zeiss MicroImaging. Reduced midbrain development was defined as >50% smaller than the mean size of controls. Each experimental group was evaluated in at least three independent experiments; a minimum of 80 individual embryos per group was examined.

**Full Methods** and any associated references are available in the online version of the paper.

**Received 11 September 2012; accepted 24 January 2013.**

**Published online 27 February 2013.**

- Sanger, H. L., Klotz, G., Riesner, D., Gross, H. J. & Kleinschmidt, A. K. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl Acad. Sci. USA* **73**, 3852–3856 (1976).
- Grabowski, P. J., Zaug, A. J. & Cech, T. R. The intervening sequence of the ribosomal RNA precursor is converted to a circular RNA in isolated nuclei of *Tetrahymena*. *Cell* **23**, 467–476 (1981).
- Danan, M., Schwartz, S., Edelheit, S. & Sorek, R. Transcriptome-wide discovery of circular RNAs in *Archaea*. *Nucleic Acids Res.* **40**, 3131–3142 (2012).
- Nigro, J. M. *et al.* Scrambled exons. *Cell* **64**, 607–613 (1991).
- Cocquerelle, C., Mascréz, B., Hetuin, D. & Bailleul, B. Mis-splicing yields circular RNA molecules. *FASEB J.* **7**, 155–160 (1993).
- Capel, B. *et al.* Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. *Cell* **73**, 1019–1030 (1993).
- Chao, C. W., Chan, D. C., Kuo, A. & Leder, P. The mouse formin (*Fmn*) gene: abundant circular RNA transcripts and gene-targeted deletion analysis. *Mol. Med.* **4**, 614–628 (1998).
- Burd, C. E. *et al.* Expression of linear and novel circular forms of an *INK4/ARF*-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet.* **6**, e1001233 (2010).
- Hansen, T. B. *et al.* miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* **30**, 4414–4422 (2011).
- Salzman, J., Gawad, C., Wang, P. L., Lacayo, N. & Brown, P. O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS ONE* **7**, e30733 (2012).
- Ambros, V. The functions of animal microRNAs. *Nature* **431**, 350–355 (2004).
- Baek, D. *et al.* The impact of microRNAs on protein output. *Nature* **455**, 64–71 (2008).
- Selbach, M. *et al.* Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58–63 (2008).
- Bartel, D. P. Circular RNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
- Krek, A. *et al.* Combinatorial microRNA target predictions. *Nature Genet.* **37**, 495–500 (2005).
- Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
- Xie, X. *et al.* Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**, 338–345 (2005).
- Friedman, R. C., Farh, K. K., Burge, C. B. & Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92–105 (2009).
- Ebert, M. S., Neilson, J. R. & Sharp, P. A. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nature Methods* **4**, 721–726 (2007).
- Franco-Zorrilla, J. M. *et al.* Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genet.* **39**, 1033–1037 (2007).
- Poliseno, L. *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038 (2010).
- Tay, Y. *et al.* Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell* **147**, 344–357 (2011).
- Cesana, M. *et al.* A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **147**, 358–369 (2011).
- Ebert, M. S. & Sharp, P. A. Emerging roles for natural microRNA sponges. *Curr. Biol.* **20**, R858–R861 (2010).
- Vivancos, A. P., Guell, M., Dohm, J. C., Serrano, L. & Himmelbauer, H. Strand-specific deep sequencing of the transcriptome. *Genome Res.* **20**, 989–999 (2010).
- Huang, R. *et al.* An RNA-Seq strategy to detect the complete coding and non-coding transcriptome including full-length imprinted macro ncRNAs. *PLoS ONE* **6**, e27288 (2011).
- Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
- Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **33**, D501–D504 (2005).
- Cabali, M. N. *et al.* Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25**, 1915–1927 (2011).
- Suzuki, H. *et al.* Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res.* **34**, e63 (2006).
- Iwai, Y., Akahane, K., Pluznik, D. H. & Cohen, R. B.  $\text{Ca}^{2+}$  ionophore A23187-dependent stabilization of granulocyte-macrophage colony-stimulating factor messenger RNA in murine thymoma EL-4 cells is mediated through two distinct regions in the 3'-untranslated region. *J. Immunol.* **150**, 4386–4394 (1993).
- Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
- Lebedeva, S. *et al.* Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR. *Mol. Cell* **43**, 340–352 (2011).
- Baltz, A. G. *et al.* The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **46**, 674–690 (2012).
- Dropcho, E. J., Chen, Y. T., Posner, J. B. & Old, L. J. Cloning of a brain protein identified by autoantibodies from a patient with paraneoplastic cerebellar degeneration. *Proc. Natl Acad. Sci. USA* **84**, 4552–4556 (1987).
- Wee, L. M., Flores-Jasso, C. F., Salomon, W. E. & Zamore, P. D. Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell* **151**, 1055–1067 (2012).
- Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnol.* **26**, 317–325 (2008).
- Landgraf, P. *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401–1414 (2007).
- Shaw, G., Morse, S., Ararat, M. & Graham, F. L. Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells. *FASEB J.* **16**, 869–871 (2002).
- Kaufman, M. H. & Bard, J. B. L. *The Anatomical Basis of Mouse Development* (Academic, 1999).
- Schambra, U. *Prenatal Mouse Brain Atlas* (Springer, 2008).
- Kapsimali, M. *et al.* MicroRNAs show a wide diversity of expression profiles in the developing and mature central nervous system. *Genome Biol.* **8**, R173 (2007).
- Jacobs, T. *et al.* Localized activation of p21-activated kinase controls neuronal polarity and morphology. *J. Neurosci.* **27**, 8604–8615 (2007).
- Chacon, M. R. *et al.* Focal adhesion kinase regulates actin nucleation and neuronal filopodia formation during axonal growth. *Development* **139**, 3200–3210 (2012).
- Kato, M. *et al.* Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
- Romeo, T. Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Mol. Microbiol.* **29**, 1321–1330 (1998).
- Gottesman, S. The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu. Rev. Microbiol.* **58**, 303–328 (2004).
- Huelsken, J., Vogel, R., Erdmann, B., Cotsarelis, G. & Birchmeier, W.  $\beta$ -Catenin controls hair follicle morphogenesis and stem cell differentiation in the skin. *Cell* **105**, 533–545 (2001).
- Park, H. C. *et al.* Analysis of upstream elements in the HuC promoter leads to the establishment of transgenic zebrafish with fluorescent neurons. *Dev. Biol.* **227**, 279–293 (2000).
- Peri, F. & Nusslein-Volhard, C. Live imaging of neuronal degradation by microglia reveals a role for v0-ATPase a1 in phagosomal fusion *in vivo*. *Cell* **133**, 916–927 (2008).
- Jeck, W. R. *et al.* Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **19**, 1–17 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank M. Feldkamp and C. Langnick (laboratory of W. Chen) for Illumina sequencing runs. We thank J. Kjems for sending us a plasmid encoding circular human CDR1 as for our zebrafish experiments. We thank K. Meier for technical assistance with zebrafish experiments and A. Sporbert from the confocal imaging facility. We thank A. Ivanov for assisting in bioinformatic analysis. N.R. thanks E. Westhof for useful discussions. We acknowledge the following funding sources: PhD program of the Max-Delbrück-Center (MDC) (S.M., F.T., L.H.G.); the MDC-NYU exchange program (M.M.); BMBF project 1210182, 'MiRNAs as therapeutic targets' (A.E.); DFG for KFO218 (U.Z.); Helmholtz Association for the 'MDC Systems Biology Network', MSBN (S.D.M.); BMBF support for the DZHK (F.I.N. and N.R.); Center for Stroke Research Berlin (J.K., F.I.N.). Funding for the group of M.L. is supported by BMBF-funding for the Berlin Institute for Medical Systems Biology (0315362C).

**Author Contributions** S.M., M.J., A.E. and F.T. contributed equally. S.M. performed many experiments, assisted by L.M. M.J. and A.E. carried out most of the computation, with contributions from N.R. and S.D.M. F.T. performed the circRNA validation experiments. A.R. performed all northern experiments. L.H.G. and M.M. contributed AGO PAR-CLIP experiments and HEK293 ribonuclease data, supervised by M.L. C.K. designed and carried out the single molecule experiments, in part together with A.L. U.Z. performed the mouse experiments. J.K. contributed the zebrafish experiments, supervised by F.I.N. N.R. designed and supervised the project. N.R. and M.J. wrote the paper.

**Author Information** Sequencing data have been deposited at GEO under accession number GSE43574. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.R. (rajewsky@mdc-berlin.de).



## METHODS

**Computational pipeline for predicting circRNAs from ribominus sequencing data.** Reference genomes (human hg19 (February 2009, GRCh37), mouse mm9 (July 2007, NB137/mm9), *C. elegans* ce6 (May 2008, WormBase v. WS190)) were downloaded from the UCSC genome browser (<http://genome.ucsc.edu/>)<sup>27</sup>. In a first step, reads that aligned contiguously and full length to the genomes were discarded. From the remaining reads we extracted 20mers from both ends and aligned them independently to find unique anchor positions within spliced exons. Anchors that aligned in the reversed orientation (head-to-tail) indicated circRNA splicing (compare main Fig. 1a). We extended the anchor alignments such that the complete read aligns and the breakpoints were flanked by GU/AG splice sites. Ambiguous breakpoints were discarded. We used the short-read mapper Bowtie 2 (ref. 52). Initially, ribominus reads were aligned in end-to-end mode to the genome:

```
$ bowtie2 -p16 --very-sensitive --phred64 --mm -M20 --score-min=C, -15, 0 -q -x <index> -U reads.qfa 2> bowtie2.log | samtools view -hbuS - | samtools sort -sample_vs_genome
```

The unmapped reads were separated and run through a custom script to split the reads as indicated in Fig. 1a to obtain 20-nucleotide anchors from both ends of the read:

```
$ samtools view -hf4 sample_vs_genome.bam | samtools view -Sb -> unmapped_sample.bam
```

```
$ ./unmapped2anchors.py unmapped_sample.bam | gzip > sample_anchors.qfa.gz
```

Here is an example of two anchor pairs in the FASTQ format; the original read was kept as part of the first anchors identifier to simplify downstream analysis:

```
@s_8_1_0001_qseq_14_A__NCCCGCCTACCGGGTCAGTGAAAAACGA
TCAGAGTAGTGGTCTTCTCCGCGGCCCGCGCGCGCGCGCTGC
NCCCGCCTACCGGGTCAGT
+
#BB@?@AB@; = @B; B@@58 (
@s_8_1_0001_qseq_14_B
CCCCGCGCGCGCGCGCTGC
+
; . : ( (>) > 0 ; . 8 #####
```

Next the anchors were aligned individually to the reference, keeping their paired ordering. The resulting alignments were read by another custom script that jointly evaluates consecutive anchor alignments belonging to the same original read, performs extensions of the anchor alignments, and collects statistics on splice sites. After the run completes, the script outputs all detected splice junctions (linear and circular) in a UCSC BED-like format with extra columns holding quality statistics, read counts etc. The original full-length reads that support each junction are written to stderr:

```
$ bowtie2 -p16 --reorder --mm -M20 --score-min=C, -15, 0 -q -x genome -U sample_anchors.qfa.gz | ./find_circ.py -S hg19 -p sample_ -s sample/sites.log > sample/sites.bed 2> sample/sites.reads
```

The resulting BED-like file is readily filtered for minimal quality cutoffs to produce the reported circRNA candidates. In particular, we demanded the following: (1) GU/AG flanking the splice sites (built in); (2) unambiguous breakpoint detection; (3) a maximum of two mismatches in the extension procedure; (4) the breakpoint cannot reside more than 2 nucleotides inside an anchor; (5) at least two independent reads (each distinct sequence only counted once per sample) support the junction; (6) unique anchor alignments with a safety margin to the next-best alignment of at least one anchor above 35 points (~more than two extra mismatches in high-quality bases); and (7) a genomic distance between the two splice sites of no more than 100 kb (only a small percentage of the data). As the ribosomal DNA cluster is part of the *C. elegans* genome assembly (ce6) and ribosomal pre-RNAs could give rise to circular RNAs by mechanisms independent

of the spliceosome, we discarded 130 candidates that mapped to the rDNA cluster on chrI:15,060,286-15,071,020.

**Permutation testing.** To test the robustness of the circRNA detection pipeline we altered the sequence of real sequencing reads in different ways at the step of anchor generation. We (1) reversed either anchor; (2) reversed the complete read; (3) randomly reassigned anchors between reads; or (4) reverse complemented the read (as a positive control). Although the reverse complement recovered the same output as expected, the various permutations led to only very few candidate predictions, well below 0.2% of the output with unpermuted reads and in excellent agreement with the results from simulated reads (Supplementary Fig. 1c).

**HEK293 RNA-seq after rRNA depletion (RibominusSeq).** Total HEK293 RNA was isolated using Trizol as recommended by the manufacturer. Ribosomal RNA was depleted from total RNA using the Ribominus kit (Invitrogen). A cDNA library was generated from rRNA-depleted RNA according to the Illumina RNA-seq protocol. The cDNA library was sequenced on an Illumina GAIIx by a 2 × 76 bp run.

***C. elegans* oocyte isolation.** Oocytes were isolated from worms carrying a temperature-sensitive (TS) allele for *fem-1* (unovulated oocytes BA17[*fem-1(hc17ts)*] strain) and *spe-9* (partially ovulated oocytes BA671[*spe-9(hc88ts)*]) as described previously<sup>53</sup>. Oocytes were washed at least four times in PBS containing protease inhibitors (Sigma-Aldrich) to separate from worm debris. Oocyte purity was observed under the dissection scope (Zeiss). Oocytes were extracted from young adults to enrich for non-endomitotic oocytes, which was also checked by fluorescence microscopy (Zeiss) with a nuclear dye. Oocytes isolated from *fem-1* or *spe-9* mutant background worms are hereafter referred to as *fem-1* oocytes and *spe-9* oocytes, respectively.

***C. elegans* sperm isolation.** Sperm was isolated in principle as described previously<sup>54</sup> from male worms obtained from a *fog-2(q71)* mutant background. Males were cut in cold PBS containing protease inhibitors (Sigma-Aldrich). Sperm was subsequently purified by filtration (3 × 40 µm nylon mesh, 2 × 10 µm nylon mesh) and a series of differential centrifugations (30 min 300g, 10 min 450g) and washed twice in cold PBS. Sperm was subsequently activated by incubation in PBS containing 200 µg ml<sup>-1</sup> Pronase (Sigma-Aldrich) for 30 min at 25 °C. Sperm purity is around 70% spermatids and spermatozoa contaminated with around 30% primary and secondary spermatocytes, as observed under oil immersion microscope.

***C. elegans* isolation of 1-cell- and 2-cell-stage embryos.** 1-cell and 2-cell-stage embryos were obtained by fluorescence-activated cell sorting as described previously<sup>55</sup>. Microscopic examination of the sorted embryos indicated that the 1-cell-stage sample was virtually pure (>98% one-cell stage embryos), whereas the 2-cell-stage embryo sample was a mixture of 1-cell-stage (40%), 2-cell-stage (55%) and older (<5%) embryos. Moreover, purity of the stages was further validated by checking for marker gene expression.

**Ribominus RNA preparation from *C. elegans* samples.** We used a kit that was developed for human and mouse samples, but still performs sufficiently to enrich mRNAs up to 30% in *C. elegans*. Most of the remaining reads mapped to ribosomal RNAs. 1 µg of total RNA per sample was depleted from rRNAs with the Ribominus Transcriptome kit (Invitrogen) according to the manufacturer's instructions with the modification that annealing of LNA probes to total RNA was performed in a thermocycler (Eppendorf) with a temperature decrease from 70 to 37 °C at a rate of 1 °C per min. Depletion of rRNAs was validated by capillary gel electrophoresis on a Bioanalyzer (Agilent). The ribominus RNA was then processed for sequencing library preparation according to the Illumina protocol.

**Cluster generation and sequencing of *C. elegans* libraries.** Cluster generation as well as sequencing of the prepared libraries was performed on the Illumina cluster station (Illumina) and sequenced on the HiSeq2000 according to the manufacturer's protocols (Illumina).

**Human gene models.** We obtained gene models for RefSeq transcripts (12 December 2011), non-coding RNAs<sup>29,56</sup> and the rnaGene and tRNA tracks from the UCSC table browser (23 April 2012)<sup>27</sup>.

**Intersection of circRNAs with known transcripts.** Our computational screen identifies only the splice sites that lead to circularization but not the internal exon/intron structure of circular RNAs. To perform analyses of the sequence content of circRNAs we therefore inferred as much as possible from annotated transcripts. The conservative assumption was that as little as possible should be spliced out. On the other hand, coincidence of circRNA splice sites with exonic boundaries inside a transcript were considered as an indicator for relevant agreement and internal introns appear to be spliced out (Supplementary Fig. 2e). We therefore sorted all overlapping transcripts hierarchically by (1) splice-site coincidence (2, 1, or 0); (2) total amount of exonic sequence between the splice sites; (3) total amount of coding sequence. The latter was used to break ties only and helped the annotation process. If one or both splice sites fell into an exon of the best matching transcript, the corresponding exon boundary was trimmed. Likewise, if it fell

into an intron or beyond transcript bounds, the closest exon was extended to match the circRNA boundaries. circRNA start/end coordinates were never altered. If no annotated exons overlapped the circRNA we assumed a single-exon circRNA. The resulting annotation of circRNAs is based on the best matching transcript and may in some cases not represent the ideal choice. Changing the annotation rules, however, did not substantially change the numbers in Fig. 1d. **Finding circRNAs conserved between human and mouse.** We reasoned that when comparing two species, the cutoff of two independent reads in each of them could be dropped, as orthologous circRNAs would automatically be supported by two independently produced reads via the intersection. We therefore mapped all mouse circRNA candidates with less stringent filtering to human genome coordinates using the UCSC liftOver tool<sup>57</sup>. The mapped mouse circRNAs were compared with independently identified human circRNAs, yielding 229 circRNAs with precisely orthologous splice sites between human and mouse. Of these, 223 were composed exclusively of coding exons and were subsequently used for our conservation analysis (Fig. 1f). When intersecting the reported sets of circRNAs supported by two independent reads in each species, we found 81 conserved circRNAs (supported by at least 4 reads in total).

**Conserved element counting.** We downloaded genome-wide human (hg19) phyloP conservation score<sup>58</sup> tracks derived from genome alignments of placental mammals from UCSC<sup>27</sup>. We interrogated the genome-wide profile inside circRNAs in two different ways. (1) Intergenic and intronic circRNAs. We read out the conservation scores along the complete circRNA and searched for blocks of at least 6-nucleotide length that exceeded a conservation score of 0.3 for intergenic and 0.5 for intronic circRNAs. The different cutoffs empirically adjust for the different background levels of conservation and were also used on the respective controls. For each circRNA, we computed the cumulative length of all such blocks and normalized it by the genomic length of the circRNA. Artefacts of constant positive conservation scores in the phyloP profile, apparently caused by missing alignment data, were removed with an entropy filter (this did not qualitatively affect the results). circRNAs annotated as intronic by the best-match procedure explained above that had any overlap with exons in alternative transcripts on either strand (five cases) were removed from the analysis. The resulting distributions are shown in Supplementary Fig. 1h, i. (2) Coding exon circRNAs. We used the best-match strategy outlined above to construct an estimated 'exon-chain' for the circRNAs that overlapped exclusively coding sequence. Using this chain we *in silico* 'spliced' out the corresponding blocks of the conservation score profile. We kept track of the frame and sorted the conservation scores into separate bins for each codon position. In addition to this, we also recorded conservation scores in the remaining pieces of coding sequence ('outside' the circRNA) as a control. However, we observed that the level of conservation is systematically different between internal parts of the coding sequence and the amino- or carboxy-terminal parts (not shown). We therefore randomly generated chains of internal exons, mimicking the exon-number distribution of real circRNAs, as a control. When analysing the circRNAs conserved between human and mouse, it became furthermore apparent that we also needed to adjust for the higher level of overall conservation. High expression generally correlates with conservation and thus, an expression cutoff was enforced on the transcripts used to generate random controls. This resulted in a good to conservative match with the actual circRNAs (Supplementary Fig. 1j, k).

**Overlap of identified circRNAs with published circular RNAs.** A number of studies in human have reported evidence for circRNAs which derive from exons of DCC<sup>4</sup>, ETS1<sup>5</sup> and a non-coding RNA from the human *INK4/ARF* locus<sup>6</sup> and the CDR1as locus<sup>9</sup>. Additionally, circRNAs from exons of the genes *CAMSAP1*, *FBXW4*, *MAN1A2*, *REXO4*, *RNF220* and *ZKSCAN1* have been recently experimentally validated<sup>10</sup>. For the four genes from the latter study, where we had ribonuclease data from the tissues in which these circRNAs were predicted (leukocytes), we recovered validated circRNAs from all of them (*ZKSCAN1*, *CAMSAP1*, *FBXW4*, *MAN1A2*).

**Cell culture and treatments.** HEK293 (Fig. 3f), HEK293TN (for virus production) and HEK293 Flp-In T-REx 293 (Life Technologies, all other experiments) were cultured in Dulbecco's modified Eagle medium GlutaMax (Gibco) 4.5 g l<sup>-1</sup> glucose, supplemented with 10% FCS, 20 U ml<sup>-1</sup> penicillin and streptomycin (Gibco) at 37 °C, 5% CO<sub>2</sub>. Whereas CDR1as/*GAPDH* ratios were within the given range, we observed two- to fivefold variation of CDR1as/vinculin ratios between different HEK lines. Transcription was blocked by adding 2 µg ml<sup>-1</sup> actinomycin D or DMSO as a control (Sigma-Aldrich) to the cell culture medium. For *in vitro* wound healing assays, cells were grown to confluency, the cell layer was disrupted using a 300 µl pipette tip and cells were washed once with medium. Bright-field images of cells were taken using a Axio Observer.Z1 (Zeiss) right after setting the scratch and 24 h later. The relative scratch areas were measured using ImageJ software.

**Quantitative PCR.** Total RNA from cell lines was isolated using Trizol (Invitrogen) extraction following the manufacturer's protocol. Adult B6129SF1/J mice

were dissected and tissue samples were collected directly into ice-cold Trizol for RNA preparation. *Caenorhabditis elegans* RNA was isolated from about 7,000 mixed stage worms by two rounds of freeze-thaw lysis in Trizol LS reagent (Invitrogen) according to the manufacturer's protocol. RNA was extracted from aqueous phase with phenol:chloroform (Ambion). RNA was precipitated with isopropanol and Glycobule (Ambion) overnight at -20 °C or for 30 min at -80 °C, respectively. Reverse transcription was performed using M-MLV (Promega) or Superscript III with oligo(dT) primer (all Invitrogen) or random primer (Metabion). For assaying mRNA expression level, qRT-PCR was performed using SYBR-Green Fluorescein (Thermo Scientific, Fermentas) and a StepOnePlus PCR System (Applied Biosystems). Expression data in CDR1as knockdown experiments, transcriptional block and RNase R assays were normalized to *C. elegans* spike-in RNA. To this end 5–10% *C. elegans* total RNA was added to the respective Trizol sample and qPCR primer for *ama-1* or *eif-3.d* were used. Mouse expression data were normalized to *Actb*. miRNA expression levels were assayed using TaqMan microRNA assays (Applied Biosystems) and normalized to sno-234. Expression levels of circRNAs described in this study were measured by qPCR using divergent primers. A list of primer sequences is available in Supplementary Table 8.

**PCR amplification and Sanger sequencing.** DNA templates were PCR amplified using BioRad Mastercycler and ThermoScientific DreamTaq Green PCR Master Mix according to the manufacturer's protocol. We performed 35 cycles of PCR. PCR products were visualized after electrophoresis in 2% ethidium bromide-stained agarose gel. To confirm the PCR results, the PCR products were purified through Agencourt AMPure XP PCR purification kit. Direct PCR product Sanger sequencing was performed by LGC Genomics Ready2 Run services. Primer P1 was provided for sequencing the product for each candidate.

**Primer design.** Divergent primers were designed for each candidate (P1, P2) to anneal at the distal ends of its sequence. As negative controls we used divergent primers for *GAPDH* and *ACTB* linear transcript in HEK293 cells, and eIF-3.D in *C. elegans*. As a further negative control for divergent primers, we used genomic DNA extracted through Qiagen DNeasy Blood & Tissue kit. As positive controls, we used convergent primers for the corresponding linear transcripts or for housekeeping genes (eIF-3.D for *C. elegans*).

**RNase R treatment.** HEK293 DNase-treated total RNA (5 µg) was incubated 15 min at 37 °C with or without 3 U µg<sup>-1</sup> of RNase R (Epicentre Biotechnologies). RNA was subsequently purified by phenol-chloroform extraction, retro-transcribed through Superscript SSIII (Invitrogen) according to the manufacturer's protocol, and used in qPCR.

**RNA nicking assay.** For partial alkaline hydrolysis (nicking) 1 µg µl<sup>-1</sup> of HEK293 total RNA was incubated in 50 mM NaHCO<sub>3</sub> for 2.5 or 5 min at 90 °C or 5 min on ice for controls. After incubation the samples were immediately resuspended in denaturing RNA sample buffer and analysed on northern blots.

**Northern blotting.** Total RNA (10–20 µg) was loaded on a 1.2% agarose gel containing 1% formaldehyde and run for 2–2.5 h in MOPS buffer.

The gel was soaked in 1×TBE for 20 min and transferred to a Hybond-N<sup>+</sup> membrane (GE Healthcare) for 1 h (15 V) using a semi-dry blotting system (Bio-Rad). Membranes were dried and ultraviolet-crosslinked (at 265 nm) 1× at 200,000 µJ cm<sup>-2</sup>. Pre-hybridization was done at 42 °C for 1 h and <sup>32</sup>P-labelled oligonucleotide DNA probes were hybridized overnight. The membranes were washed briefly in 2× SSC, 0.1% SDS at room temperature and two additional times at 55 °C for 30 min, followed by two 30-min washes in 0.2× SSC, 0.1% SDS at 50–55 °C. For data collection, the membrane was exposed to a phosphorimager screen.

**Genome alignments for detecting miRNA seed complementary sites.** Multiple species alignments for the genomic intervals, corresponding to circRNAs predicted in *C. elegans* (ce6), human (hg19) and mouse (mm9), were generated via the Galaxy server at UCSC<sup>59–61</sup>. In case that a circle was overlapping with an annotated transcript, the inferred spliced sequence was used for retrieving the alignments.

The alignments included *C. elegans*, *C. briggsae* and *C. remanei* in the first case and *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Bos taurus* and *Canis familiaris* in the latter two.

**C. elegans human and mouse miRNAs.** Fasta files with *C. elegans*, human and mouse miRNAs were obtained from miRBase release 16 (ref. 62). Only mature miRNAs were considered for the seed analysis. According to miRBase 16 a mature miRNA is the predominant miRNA between the two species arising from the two arms of the precursor hairpin (information that is not included in more recent versions). The miRNAs were grouped into families that share a common seed (nucleotides 2–7). There are 117, 751 and 723 miRNA families for *C. elegans*, human and mouse, respectively.

**Detecting putative miRNA seed matches.** The *C. elegans*, human and mouse multiple species alignments were scanned for putative conserved miRNA target sites for each of the mature miRNA families. A putative target site of a miRNA is a



6-nucleotide-long sequence in the genome that is the reverse complement of nucleotides 2–7 of the mature miRNA sequence. A putative target site is called conserved if it is found in *C. elegans*, *C. briggsae* and *C. remanei* in the first case or in human, mouse, rat, cow and dog in the latter.

**AGO PAR-CLIP.** Generation and growth conditions of human embryonic kidney (HEK) 293 cells and HEK293 stably expressing Flag/HA-AGO1 and Flag/HA-AGO2 were reported previously<sup>63</sup>. Stably transfected and parental HEK293 cells were labelled with 100  $\mu$ M 4-thiouridine for 16 h. After labelling, procedure followed the PAR-CLIP protocol as described<sup>32</sup>. Briefly, ultraviolet-irradiated cells were lysed in NP-40 lysis buffer. Immunoprecipitation was carried out with protein G magnetic beads (Invitrogen) coupled to anti-Flag antibody (Sigma) and to anti-AGO2 antibody<sup>64</sup> from extracts of stably transfected and parental HEK293 cells, respectively, for 1 h at 4 °C. Beads were treated with calf intestinal phosphatase (NEB) and radioactively end-labelled by T4 polynucleotide kinase (Fermentas). The crosslinked protein–RNA complexes were resolved on 4–12% NuPAGE gel (Invitrogen), and a labelled protein–RNA complex of close to 100 kDa was excised. The protein–RNA was isolated by electroelution. RNA was isolated by proteinase K treatment and phenol-chloroform extraction, reverse transcribed and PCR-amplified. The amplified cDNA was sequenced on a GAIIX (Illumina) with 36 cycles.

**Human Argonaute PAR-CLIP analysis.** We obtained Argonaute PAR-CLIP reads from ref. 32. We additionally produced 4 PAR-CLIP libraries. In total, we analysed the following PAR-CLIP data sets: AGO1\_4su\_1 (SRR048973), AGO3\_4su\_1 (SRR048976) from ref. 32; AGO1\_4su\_ML\_MM\_6, AGO2\_4su\_ML\_MM\_7, AGO2\_4su\_ML\_MM\_8, and AGO2\_4su\_3\_ML\_LG (our own data, published under GEO accession GSE43574).

Redundant reads were collapsed (such that each distinct read sequence appears only once), aligned to the human genome (assembly hg19) using bwa 0.6.1-r104 (ref. 65), and analysed by our in-house PAR-CLIP analysis pipeline (Jens, M. *et al.*, unpublished), essentially as described in ref. 33. Briefly, reads uniquely aligning to the genome are grouped into clusters contiguously covering the reference, assigning each cluster a number of quality scores (T conversions, number of independent reads, etc.). Clusters with less than 3 reads from 3 of 6 independent AGO PAR-CLIP libraries or lacking T conversions were discarded. Remaining clusters are annotated against a comprehensive list of transcript models (see below) and collected into ‘only sense’, ‘only antisense’ and ‘intergenic/overlapping transcription’ categories based on their annotation. As PAR-CLIP sequencing preserves the directionality of RNA fragments we assume ‘only antisense’ clusters to predominantly represent false positives due to mapping artefacts (PAR-CLIP RNA is mutated and fragments are often short), and choose quality cutoffs for all clusters such that the fraction of kept ‘only antisense’ clusters is reduced to below 5%. Remaining ‘only antisense’ clusters were discarded. For Fig. 3a, uniquely aligning, collapsed reads are shown.

**AGO binding sites in *C. elegans*.** Sequencing reads from the Zisoulis Alg-1 HITS-Clip data were obtained from [http://yeolab.ucsd.edu/yeolab/Papers\\_files/ALG1\\_MT-WT\\_raw.tar.gz](http://yeolab.ucsd.edu/yeolab/Papers_files/ALG1_MT-WT_raw.tar.gz) (ref. 66). The raw sequencing data of the wild-type Alg-1 HITS-CLIP was pre-processed and mapped with the mapper module from miRDeep2 (ref. 74). The pre-processed reads were mapped with bowtie version 0.12.7 (ref. 67) to the *C. elegans* genome (ce6). All reads that overlapped when mapped to the genome were merged into bigger regions (islands). Read counts were averaged. This resulted in 24,910 islands in the *C. elegans* genome.

**Analysis of sequence conservation in CDR1as.** Genome alignments of 32 vertebrates were downloaded from the UCSC database (hg19)<sup>27</sup> and analysed for the CDR1as locus. Primate species other than human were discarded to not bias the analyses. The one species (cow) with more than 50% gaps in the CDR1as locus was also discarded. The alignments for the seed regions were then corrected. Specifically, bases that would clearly align with the seed but had been separated in the alignment by runs of gaps were re-aligned. These corrections were necessary in less than 1% of all seed sites.

For an in-depth analysis we BLATed<sup>68</sup> the human CDR1as sequence with 20-nucleotide flanking region against all vertebrate genomes in the UCSC genome browser and kept only hits that in turn aligned best to the human locus. The resulting sequences were used to build a multiple species alignment with MUSCLE<sup>69</sup>. The same corrections were applied as described above. This alignment was also used for Supplementary Fig. 4. Entropy was calculated in log<sub>2</sub> units and averaged across all alignment columns bracketing each human seed site by maximally 8 nucleotides.

**Analysis of miR-7 base-pairing within CDR1as.** RNAfold<sup>70</sup> was used to co-fold miR-7 with each of the 74 binding regions within CDR1as defined as the miR-7 seed match TCTTCC and the next 16 bases upstream.

**Single-molecule RNA fluorescence in situ hybridization (smRNA FISH).** 48 oligonucleotide probes (20 nucleotides length; spacing 2 nucleotides) complementary to the CDR1as transcript were designed using the Stellaris Probe

Designer version 2.0 (Biosearch Technologies) with a masking level of 4 on the human genome to achieve high probe specificity (Supplementary Table 8). Stellaris probe pools were obtained from BioCat GmbH as conjugates coupled to Quasar 670 (a Cy 5 replacement). Flp-In T-REX 293 cells (Life Technologies) were grown exponentially and seeded into LabTek 4-well chambered coverslips (1 to  $2 \times 10^5$  cells per well). Hybridizations were performed according to the manufacturer's instructions with 50 ng ml<sup>-1</sup> DAPI as nuclear counterstain; Stellaris probes were hybridized at 125 nM concentration with a stringency of 10% formamide in overnight hybridizations at 37 °C. Images were acquired on an inverted Nikon Ti microscope with a Hamatsu ORCA R2 CCD camera, a 60 $\times$  NA 1.4 oil objective and Nikon NIS-Elements Ar software (version 4), using an exposure time of 50 ms for DAPI and 1–1.5 s for Quasar 670. Groups of cells for imaging were chosen in the DAPI channel; Z-stacks were acquired in the Quasar 670 channel using 0.3  $\mu$ m spacing and comprised a total depth of 6.5  $\mu$ m (5  $\mu$ m below and 1.5  $\mu$ m above the middle of the nucleus) and merged using maximum intensity.

**Mouse strains and in situ hybridizations.** All mice were bred and maintained in the animal facility of the Max Delbrück Centrum under specific pathogen-free conditions, in plastic cages with regular chow and water ad libitum. All aspects of animal care and experimental protocols were approved by the Berlin Animal Review Board (REG 0441/09). B6129SF1/J wild-type adult, newborns (postnatal day 1) or pregnant females (plug detection at day 0.5; embryo collection at day 13.5) were used, as indicated for each experiment, to obtain the tissues needed for RNA analysis and in situ hybridizations (ISH). After death, embryos or tissues were immediately frozen in liquid nitrogen and stored at –70 °C, or fixed for ISH.

Mouse brain structures were collected and named according to the anatomical guidelines of the Gene Expression Nervous System Atlas of the Rockefeller University (<http://www.gensat.org>) and the Mouse Brain Atlas ([http://www.mbl.org/mbl\\_main/atlas.html](http://www.mbl.org/mbl_main/atlas.html)).

For the RNA analysis and to clone CDR1as-specific RNA probes, two adult 1-year-old mice of both sexes were dissected, total RNA prepared and analysed. If embryos or newborns were sectioned, a minimum of two specimens were evaluated; in some instances up to 5 specimens were used.

For ISH, samples were fixed in formalin (1 $\times$ PBS; 4% formaldehyde) for 12 h and post-fixed (70% ethanol, 18 h) before dehydrating and paraffin-embedding. Next, the organs were perfused with a standard protocol using a Shandon XP Hypercentre. For ISH mouse embryos or organs were cut in RNase-free conditions at 6  $\mu$ m and ISH was performed as described<sup>48</sup> with digoxigenin (DIG)-labelled RNA probes. All DIG–RNA probes were hybridized at 58 °C overnight. A total of 600 ng of the labelled probes was used per slide.

To amplify Cdr1 sense and antisense sequences for ISH probe preparation a standard PCR-amplification was performed using mouse cerebellum cDNA. Three Cdr1 amplicons were generated, two of which probes are meant for the detection of both linear and circular forms using mmuCdr1\_1f 5'-TGCCAGTACCAAGGTCTTCC-3' and mmuCdr1\_1r 5'-TTTTCTGCTGGAAGATGTCAA-3', as well as mmuCdr1\_2f 5'-CCAGACAATCGTGATCTTCC-3' and mmuCdr1\_2r 5'-ATCTTGCGCTGGAAGACTTGG-3'. In addition a probe was generated, specific to the circular probe, using the divergent primers mmuCdr1\_as\_7f 5'-CCACATCTTCCAGCATCTTT-3' and mmuCdr1\_as\_7r 5'-TGGATCCCTTGAAGACAAA-3' (CDR1as head to tail probe). All ensuing fragments were subcloned into pCR-BluntII-TOPO (Invitrogen) and verified by sequencing. Linearized plasmids were amenable for in vitro transcription using the T7 (antisense) or SP6 (sense) polymerase and a DIG-label nucleotide mixture according to manufacturer's instruction (Roche Applied Science).

LNA ISHs were performed according to a protocol suggested by the manufacturer (Exiqon) with minor modifications. For individual LNAs, specific protocols were run at 51 °C (miR-7; 38485-15) or 58 °C (miR-124; 88066-15) on an InSituPro VS robot (Intavis). A pre-hybridization step was added, which consisted of an incubation of the slides at 15 °C lower than the hybridization temperature for 30 min using hybridization buffer. The antibody-blocking step was performed in the presence of 1% mouse blocking reagent (Roche 11096176001) and 10% sheep serum. The LNA probes were used at the following concentrations: miR-7 40 nM; miR-124 20 nM; U6 snRNA 1 nM; scrambled 40 nM, as suggested by miRCURY LNA microRNA ISH Optimization kit (Exiqon; 90004). Before detection all slides were washed 4 $\times$  in NTMT including 1 mM Levamisole. The doubly DIG-labelled LNAs were detected by the alkaline phosphatase using the substrate BM-purple (Roche; 11442074001) at 37 °C.

**siRNA- and shRNA-mediated knock down.** CDR1as was knocked down using custom designed siRNA oligonucleotides (Sigma) and Lipofectamine RNAiMax (Invitrogen).  $2 \times 10^6$  HEK293 cells were transfected with 10 nM siRNA duplex following the manufacturer's protocol. After 12–16 h cells were harvested and subjected to RNA analysis. For stable knock down of CDR1as, 293TN cells were co-transfected with the packaging plasmids pLP1, pLP2 and the VSV-G plasmid

(Invitrogen) and pSicoR constructs<sup>71</sup> (sequences available in the Supplementary Table 8) by calcium phosphate transfections. Viral supernatants were harvested after 24 h and 48 h post transfection and filtered through a 0.44 µm filter. For infection the viral supernatants supplemented with fresh medium and 6 µg ml<sup>-1</sup> polybrene was added to target cells. After overnight infection cells were allowed to recover for 12 h and subjected to a second round of infection. Cells were collected 48–72 h after the first infection. The list of siRNA oligonucleotides is provided in Supplementary Table 8.

**Zebrafish methods.** Zebrafish and their embryos were handled according to standard protocols<sup>72</sup> and in accordance with Max Delbrück Centrum institutional ethical guidelines. The *Tg(huC:egfp)* and the *Tg(Xia.Tubb:dsRED)* transgenic zebrafish lines have been described elsewhere<sup>49,50</sup>. Morpholino antisense oligomers (Gene Tools) were prepared at a stock concentration of 1 mM according to the manufacturer's protocol. Sequences: control morpholino, 5'-CTC TTACCTCAGTTACAATTTATA-3' (control morpholino) and morpholino targeting miR-7, 5'-ACAACAAAATCACAAGTCTTCCACA-3' (miR-7 morpholino). For titration experiments we used 15 ng of control morpholino and 9 and 15 ng of miR-7 morpholino; for all other experiments we used 9 ng miR-7 morpholino. 3 nl of morpholinos were injected into the yolk of single-cell-stage embryos.

A 673-nucleotide mouse *Cdr1as* fragment was amplified from mouse cerebellar cDNA and the amplicon was subcloned into a pCR-Blunt II Topo vector (Invitrogen). The vector was linearized with KpnI or ApaI (Fermentas) *in vitro* transcribed (IVT) using T7 and SP6 RNA polymerases (Promega) and the resulting *Cdr1as* and reverse complement *Cdr1as* control products were used for injections (1.5 nl of 100 ng nl<sup>-1</sup>) into the cell of single-cell-stage embryos. In a repetition of these experiments the *Cdr1as* fragment amplicon was directly used as a template for IVT by exploiting T7-promoter extended forward and reverse primer.

Approximately 1.5 nl of a 50 ng µl<sup>-1</sup> construct (backbone pCS2+) expressing the human linear or the human circular CDR1as was injected into the cell of single-cell-stage embryos (provided by the Kjems laboratory). For rescue experiments the construct containing the circular CDR1as was injected together with 1.5 nl pre-miR-7 precursor (7 µM, pre-miR miRNA precursor ID PM10047 from Applied Biosystems). The negative control was the vector pCS2+ without insert (empty vector, 50 ng µl<sup>-1</sup>).

Confocal imaging was performed using a Zeiss LSM 510 microscope (Carl Zeiss MicroImaging) equipped with a 25× objective (NA = 0.8). Embryos were anaesthetized using 0.1% tricaine and mounted in 1% agarose as described<sup>73</sup>. Confocal stacks were acquired of the brain region and processed using Zeiss ZEN software. Midbrain and telencephalon volumes were calculated using Imaris 64×7.6.1 software based on high-resolution three-dimensional stacks obtained from *Tg(Xia.Tubb:dsRED)* embryos. Reduced midbrain development was defined as >50% smaller than the mean size of controls.

Each experimental group was evaluated in at least three independent experiments; a minimum of 80 individual embryos per group were examined. Data are expressed as mean ± standard deviation. Statistical analysis was performed using Student's *t*-test, and a *P* < 0.05 was considered statistically significant.

Expression of miR-7 in zebrafish embryos at 48 hours post fertilization was normalized to expression of β-actin. In the miR-7 morpholino group, only embryos with a midbrain phenotype were used for the RNA expression analysis.

dre β-actin forward primer, 5'-TGCTGTTTTCCCTCCATTG-3'; reverse primer, 5'-TTCTGTCCCATGCCAACCA-3'; probe sequence FAM-5'-TGGAC GACCCAGACATCAGGGAGTG-3'-TAMRA.

For measuring the expression of dre-miR-7a/b we used Applied Biosystems TaqMan miR assays (ID000268, ID001088).

52. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
53. Aroian, R. V., Field, C., Pruliere, G., Kenyon, C. & Alberts, B. M. Isolation of actin-associated proteins from *Caenorhabditis elegans* oocytes and their localization in the early embryo. *EMBO J.* **16**, 1541–1549 (1997).
54. L'Hernault, S. W. & Roberts, T. M. Cell biology of nematode sperm. *Methods Cell Biol.* **48**, 273–301 (1995).
55. Stoeckius, M. *et al.* Large-scale sorting of *C. elegans* embryos reveals the dynamics of small RNA expression. *Nature Methods* **6**, 745–751 (2009).
56. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnol.* **28**, 511–515 (2010).
57. Hinrichs, A. S. *et al.* The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* **34**, D590–D598 (2006).
58. Pollard, K. S. *et al.* Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20**, 110–121 (2010).
59. Blankenberg, D. *et al.* Galaxy: a web-based genome analysis tool for experimentalists. *Curr. Protoc. Mol. Biol.* Ch. 19, Unit 19 10 11–21 (2010).
60. Giardine, B. *et al.* Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* **15**, 1451–1455 (2005).
61. Goecks, J., Nekrutenko, A., Taylor, J. & Galaxy, T. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11**, R86 (2010).
62. Griffiths-Jones, S. The microRNA Registry. *Nucleic Acids Res.* **32**, D109–D111 (2004).
63. Landthaler, M. *et al.* Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA* **14**, 2580–2596 (2008).
64. Rudel, S., Flatley, A., Weinmann, L., Kremmer, E. & Meister, G. A multifunctional human Argonaute2-specific monoclonal antibody. *RNA* **14**, 1244–1253 (2008).
65. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
66. Zisoulis, D. G. *et al.* Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nature Struct. Mol. Biol.* **17**, 173–179 (2010).
67. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
68. Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
69. Goujon, M. *et al.* A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* **38**, W695–W699 (2010).
70. Bernhart, S. H. *et al.* Partition function and base pairing probabilities of RNA heterodimers. *Algorithms Mol. Biol.* **1**, 3 (2006).
71. Ventura, A. *et al.* Cre-lox-regulated conditional RNA interference from transgenes. *Proc. Natl Acad. Sci. USA* **101**, 10380–10385 (2004).
72. Westerfield, M. *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Brachydanio rerio)* 2nd edn (Univ. Oregon Press, 1993).
73. Krueger, J. *et al.* Flt1 acts as a negative regulator of tip cell formation and branching morphogenesis in the zebrafish embryo. *Development* **138**, 2111–2120 (2011).
74. Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **40**, 37–52 (2012).



# ATPase-dependent quality control of DNA replication origin licensing

Jordi Frigola<sup>1</sup>, Dirk Remus<sup>1†</sup>, Amina Mehanna<sup>1</sup> & John F. X. Diffley<sup>1</sup>

**The regulated loading of the Mcm2–7 DNA helicase (comprising six related subunits, Mcm2 to Mcm7) into pre-replicative complexes at multiple replication origins ensures precise once per cell cycle replication in eukaryotic cells. The origin recognition complex (ORC), Cdc6 and Cdt1 load Mcm2–7 into a double hexamer bound around duplex DNA in an ATP-dependent reaction, but the molecular mechanism of this origin ‘licensing’ is still poorly understood. Here we show that both Mcm2–7 hexamers in *Saccharomyces cerevisiae* are recruited to origins by an essential, conserved carboxy-terminal domain of Mcm3 that interacts with and stimulates the ATPase activity of ORC–Cdc6. ATP hydrolysis can promote Mcm2–7 loading, but can also promote Mcm2–7 release if components are missing or if ORC has been inactivated by cyclin-dependent kinase phosphorylation. Our work provides new insights into how origins are licensed and reveals a novel ATPase-dependent mechanism contributing to precise once per cell cycle replication.**

Stable genome inheritance requires that replication origins initiate efficiently during S phase, and that re-initiation from these origins is subsequently prevented. This is accomplished by first licensing origins during G1 phase with a pre-replicative complex containing an inactive double hexamer of the Mcm2–7 helicase, and then activating the helicase during S phase, when pre-replicative complexes can no longer be assembled<sup>1–9</sup>.

## Mcm3 recruits Cdt1–Mcm2–7 to ORC–Cdc6

We have previously described a system in which pre-replicative complexes can be assembled on DNA with purified proteins<sup>7</sup>. When ATP is present, ORC and Cdc6 load Cdt1–Mcm2–7 heptamers into a salt-resistant Mcm2–7 double hexamer with concomitant release of Cdc6 and Cdt1; in ATP- $\gamma$ S (a slowly hydrolysed analogue of ATP) all pre-replicative complex components are recruited in low salt but are removed by a high-salt wash (HSW) (Supplementary Fig. 1a). To investigate the individual roles of Cdt1 and Mcm2–7, we purified each separately (Supplementary Fig. 1b–d) and showed each protein was functional for loading when re-assembled into a Cdt1–Mcm2–7 complex (Supplementary Fig. 2). Consistent with our previous results<sup>7</sup>, none of the pre-replicative complex components bound to DNA under recruitment conditions (ATP- $\gamma$ S, low salt) in the absence of ORC (Fig. 1A, lanes 1–6). In the absence of Cdc6 (lanes 7–9), ORC bound DNA, but could not recruit Cdt1, Mcm2–7 or Cdt1–Mcm2–7. Cdc6 was recruited to DNA in an ORC-dependent manner (compare lanes 4 and 10) consistent with the formation of an essential ORC–Cdc6 complex<sup>10</sup>. Cdt1 was not recruited in the presence of ORC–Cdc6 (lane 10), but we found significant ORC- and Cdc6-dependent recruitment of Mcm2–7 subunits in the absence of Cdt1 (compare lane 11 to lanes 6 and 9). Cdt1 was only recruited when Mcm2–7 was present along with ORC and Cdc6 (lane 12). From this we conclude that Mcm2–7 can be recruited to ORC–Cdc6 without Cdt1 and that Cdt1 recruitment requires Mcm2–7.

The absence of Cdt1 did not affect Mcm7 recruitment, but did reduce the amount of Mcm2 recruited (Fig. 1A, compare lanes 11 and 12). To explore this further, we tested for the presence of each of the six Mcm2–7 subunits after Mcm2–7 recruitment in ATP- $\gamma$ S in

the presence or absence of Cdt1 using silver staining (Fig. 1B, a) and immunoblotting (Fig. 1B, b). We found that Mcm3, 5 and 7 were recruited to ORC–Cdc6 similarly in the presence and absence of Cdt1 (Fig. 1B); the recruitment of Mcm2, 4 and 6, however, was significantly reduced without Cdt1. This suggests that Mcm3, 5 and 7, which are immediate neighbours in the Mcm2–7 ring, are recruited to ORC–Cdc6 directly without Cdt1, but Cdt1 plays some role in recruiting Mcm2, 4 and 6 (Fig. 1B, c).

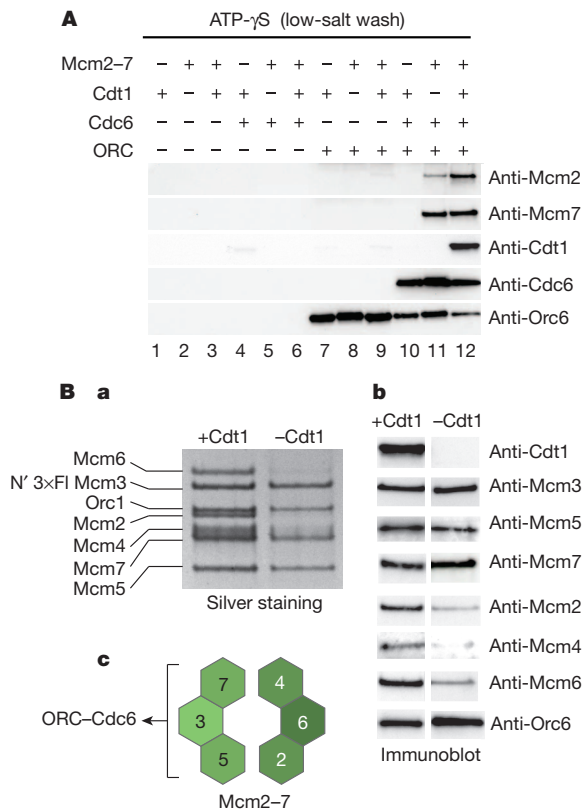
## The extreme C terminus of Mcm3 recruits Cdt1–Mcm2–7

We took several approaches to address which Mcm2–7 subunits were involved in the Cdt1-independent recruitment. We tested each of the six Mcm2–7 subunits individually and found that only Mcm3 could be recruited in a Cdc6-dependent manner without the other subunits (Supplementary Fig. 3a, b, f). We next assembled Mcm2–7 complexes containing all subunit or lacking either Mcm3 or Mcm4 ( $\Delta 3$  and  $\Delta 4$ , respectively; Supplementary Fig. 3c, d). Mcm5 and 7 were recruited to ORC–Cdc6 in the absence of Cdt1 by the complete complex (Supplementary Fig. 3f, lanes 1 and 5) and by  $\Delta 4$  (lanes 3 and 7), but not by  $\Delta 3$  (lanes 2 and 6). Taken together, these results indicate that Mcm3 is critical for recruiting Mcm5 and 7 to ORC–Cdc6 in the absence of Cdt1.

Mcm3 comprises an amino-terminal domain (NTD) and a AAA+ domain, which are both found in all Mcm2–7 subunits<sup>11,12</sup>, and an extended C-terminal tail of unknown function containing a conserved domain at the extreme C terminus (‘C’ in Fig. 2a) that is not found in other Mcm2–7 subunits (Fig. 2a). A version of Mcm3 containing a 3 $\times$ Flag tag at its C terminus could be assembled into a stable heptameric Cdt1–Mcm2–7 complex, but was unable to recruit or load Mcm2–7 (Supplementary Fig. 4). Moreover, a small fragment containing the C-terminal 194 amino acids residues of Mcm3 was recruited in an ORC- and Cdc6-dependent manner (Fig. 2b). These results pointed to the C terminus as being crucial for Mcm3 recruitment.

Based on the conservation of the Mcm3 C terminus (Fig. 2a), and the fact that a tag at the C terminus interfered with its function, we generated a series of C domain amino acid substitution mutants in full-length Mcm3 (Fig. 2a). We assembled the mutant proteins into

<sup>1</sup>Cancer Research UK London Research Institute, Clare Hall Laboratories, South Mimms EN6 3LD, UK. <sup>†</sup>Present address: Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10065, USA.



**Figure 1 | Mcm3 is necessary and sufficient for Mcm2-7 recruitment.** **A**, Protein requirements for Mcm2-7 recruitment. ORC, Cdc6, Cdt1 and Mcm2-7 were purified as described in Methods and in Supplementary Fig. 1. After incubation of the indicated proteins with DNA beads in the presence of ATP-γS, beads were isolated and washed with low salt. DNA was uncoupled with irradiation as described in Methods and bound proteins were analysed by immunoblot with the indicated antibodies. **B**, Recruitment of Mcm2-7 by ORC and Cdc6 was performed as in **A**. Reactions contained Mcm2-7 either with (+) or without (-) Cdt1. **B**, **a**–**c**, Bound proteins were visualized by silver staining (**B**, **a**) or immunoblot (**B**, **b**). Panel **C**, **c** summarizes the results in **B**, **a**, **b**.

Cdt1-Mcm2-7 complexes (Supplementary Fig. 5) and tested their ability to recruit and load Mcm2-7. The single mutants Mcm3-11 and Mcm3-12, and the Mcm3-13 double mutant were completely defective in recruitment of Cdt1-Mcm2-7 to ORC-Cdc6 (Fig. 2c, lanes 2–4), as was Mcm3 lacking its entire C-terminal domain (3ΔC, Fig. 2c, lane 8). In these experiments, Orc1 serves as a useful loading control. These mutant proteins were also not recruited when tested individually in the absence of Cdt1 and the other Mcm2-7 subunits (Supplementary Fig. 6). Consistent with their defect in recruitment, Mcm3-11 and Mcm3ΔC were unable to load Mcm2-7 (Fig. 2d, lanes 2 and 6). Mutation of other conserved residues (Mcm3-16, Mcm3-22 and Mcm3-23) generated proteins that were less defective, showing reduced recruitment (Fig. 2c, lanes 5–7 and Supplementary Fig. 6) and a commensurate reduction in loading (Fig. 2d, lanes 3–5). Therefore, the C terminus of Mcm3 is essential for recruitment of all Mcm2-7 subunits. As Mcm2-7 containing Mcm3-11, Mcm3-12 or Mcm3-13 is not recruited even when bound to Cdt1, this indicates that Cdt1 cannot recruit Mcm2-7 to ORC-Cdc6 directly. The reduced association of Mcm2, 4, 6 with ORC-Cdc6 (Fig. 1A, B) therefore probably reflects a role for Cdt1 in retaining or stabilizing these subunits after initial recruitment by Mcm3.

We next expressed the mutant Mcm3 proteins in yeast strains harbouring a temperature-sensitive degron mutant of *MCM3* (*mcm3-ta*)<sup>13</sup>. This degron mutant grew well at 25 °C, but not at the restrictive temperature, 37 °C (Fig. 2e, row 1). The growth defect at 37 °C was suppressed by expression of wild-type Mcm3 (Mcm3,

row 2), but not by expression of Mcm3-11, Mcm3-12 or Mcm3-13 (rows 3–5) indicating that these proteins are not functional *in vivo*. Expression of these proteins significantly reduced growth even at 25 °C when the wild-type degron fusion protein was functional (middle panel, GAL 25 °C, compare row 1 with rows 3–5), indicating that these mutants act as dominant-negative inhibitors when over-expressed. Some of the other mutants (Mcm3-16, Mcm3-22, Mcm3-23) showed mild growth defects at 37 °C, consistent with their partial defects in Mcm2-7 loading. Supplementary Fig. 7 shows that *mcm3-11* supported only very slow growth when present as an only copy, whereas *mcm3-12* and *mcm3-13* did not support growth. Taken together, these results indicate that a domain at the extreme C terminus of Mcm3 is necessary for recruiting Cdt1/Mcm2-7 to ORC-Cdc6, and that this function is essential for viability.

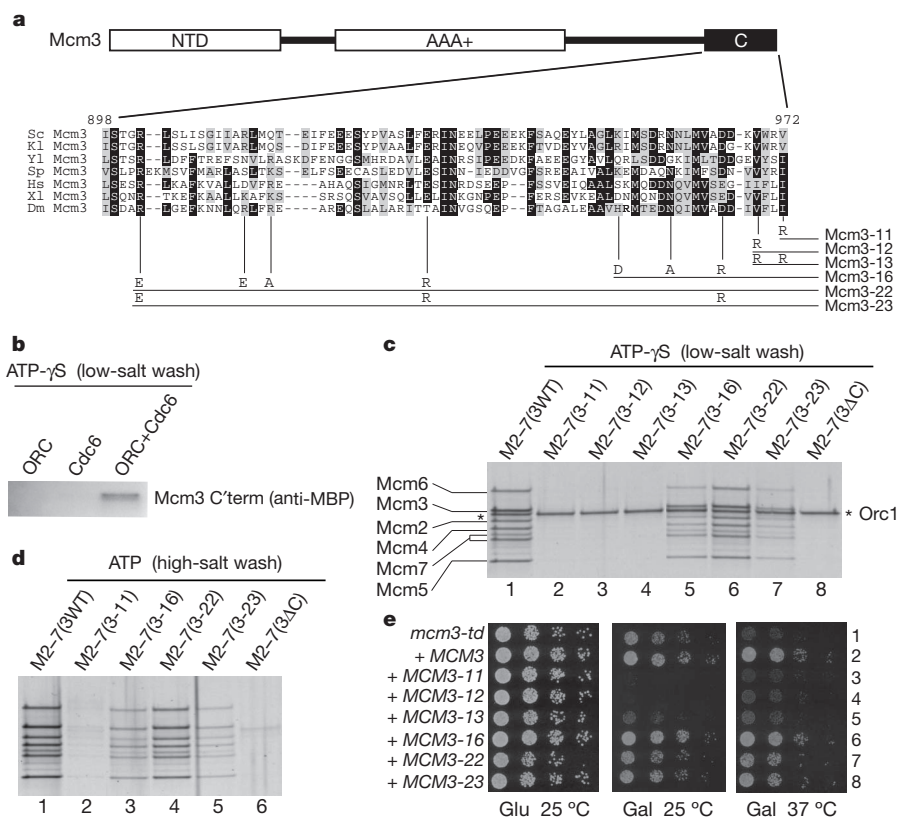
## Both Mcm2-7 hexamers require the C terminus of Mcm3

The results in Figs 1 and 2 show that recruitment of the first Mcm2-7 hexamer occurs by interaction of Mcm3 with ORC-Cdc6. If recruitment of the second hexamer occurred by a different mechanism, one that didn't involve Mcm3-ORC-Cdc6 interaction but did require recruitment of the first hexamer, then the Mcm3-11 mutant should be able to be recruited as the second hexamer in the presence of wild-type Cdt1-Mcm2-7. To test this, we first generated two different forms of Mcm3 that could be distinguished after SDS-polyacrylamide gel electrophoresis (SDS-PAGE), one containing an N-terminal fusion to the Flag tag (N' 3×Flag-Mcm3), the other containing a larger N-terminal fusion to maltose binding protein (N' MBP-Mcm3). Both proteins supported similar levels of recruitment and loading, and both complemented the *mcm3-ta* mutant at 37 °C (Supplementary Fig. 8). We engineered the Mcm3-11 mutation into either the MBP fusion (Fig. 3a) or the Flag fusion (Fig. 3b), generated mutant Cdt1-Mcm2-7 complexes and mixed these complexes in different ratios with Cdt1-Mcm2-7 containing either wild-type Flag-Mcm3 (Fig. 3a) or wild-type MBP-Mcm3WT (Fig. 3b). These experiments show that the Mcm3-11 mutant was not recruited above background levels in ATP or ATP-γS over a wide range of wild type:mutant ratios, showing that the wild-type protein cannot aid recruitment of the Mcm3-11 mutant complex, indicating that recruitment of both hexamers requires the Mcm3 C terminus.

## Mcm3 binding activates the ORC-Cdc6 ATPase

ATP hydrolysis is required for assembly of the Mcm2-7 double hexamer, but what stage of assembly requires ATP hydrolysis is not known. We therefore tested whether Mcm3 binding affected the ATPase activity of DNA-bound ORC-Cdc6. Neither Mcm3 nor the Mcm3-13 mutant showed appreciable ATPase activity by themselves (Fig. 4a), consistent with the fact that, as a AAA+ ATPase, Mcm3 requires an arginine finger from a binding partner (in this case, Mcm5) to promote ATP hydrolysis. ORC, together with Cdc6 (ORC-Cdc6, Fig. 4a), had some ATPase activity in the presence of DNA, which might reflect a basal level of ATPase from DNA-bound ORC-Cdc6, or might be caused by ORC and/or Cdc6 that is not bound to DNA. Addition of Mcm3, but not Mcm3-13, to ORC-Cdc6 resulted in a three- to fourfold increase in ATPase activity, indicating that the interaction of the C terminus of Mcm3 stimulates the ATPase of ORC-Cdc6. To rule out any contribution from the Mcm3 ATPase, we examined the effect of the Mcm3 C-terminal fragment lacking the AAA+ domain. Figure 4b shows that this domain was as effective as full-length Mcm3 in stimulating the ATPase activity of ORC-Cdc6. Reactions in Fig. 4b contained 2.5 pmol each of ORC and Cdc6, yet addition of Mcm3 resulted in approximately 140 pmol of additional ATP hydrolysis (above the level produced by ORC and Cdc6 alone) over 20 min, indicating that Mcm3 binding can induce multiple rounds of ATP binding and hydrolysis by ORC-Cdc6.





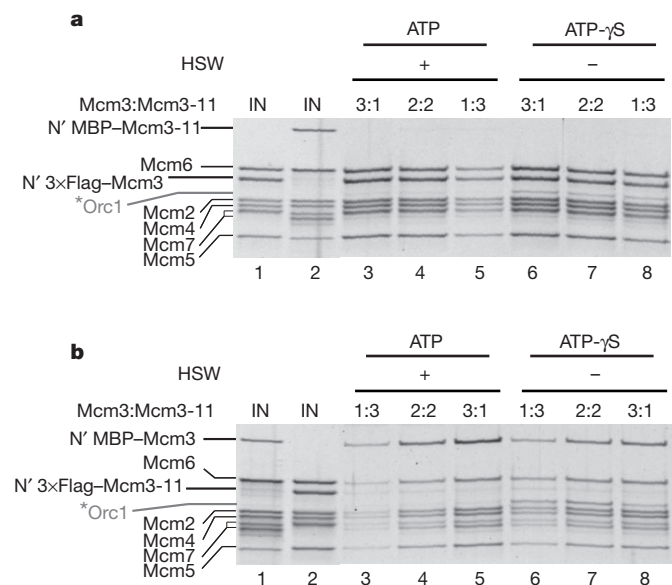
**Figure 2 | The C terminus of Mcm3 is required for Mcm2-7 recruitment.** **a**, The domain architecture of Mcm3 and alignment of the Mcm3 C terminus. The alignment includes Mcm3 from a variety of eukaryotic species (Sc, *Saccharomyces cerevisiae*; Kl, *Kluyveromyces lactis*; Yl, *Yarrowia lipolytica*; Sp, *Schizosaccharomyces pombe*; Hs, *Homo sapiens*; Xl, *Xenopus laevis*; Dm, *Drosophila melanogaster*). Residue numbers above alignment correspond to *S. cerevisiae* Mcm3. The position of various mutants is shown by vertical lines, and the mutant amino acid residue is shown at the bottom of the line. Allele names are on the right. The numbering of alleles begins with 11 to prevent confusion with any existing *mcm3* mutant alleles. **b**, Recruitment assay performed with a C-terminal MBP-tagged Mcm3 fragment (last 194 amino acids) and with the indicated proteins. **c**, **d**, Cdt1-Mcm2-7 complexes (M2-7) containing wild-type (WT) and mutant Mcm3 (digits shown in parentheses) were tested for recruitment (**c**) or for loading (**d**), proteins were analysed by silver staining. **e**, Wild-type and mutant *mcm3* alleles were expressed from a galactose-inducible promoter in strains containing the *mcm3*-td degon mutant. A dilution series was tested for growth on plates containing either 2% D-glucose (Glu) or 2% D-galactose (Gal) at either 25 °C or 37 °C as indicated.

### ATP hydrolysis by ORC-Cdc6 promotes Mcm2-7 release

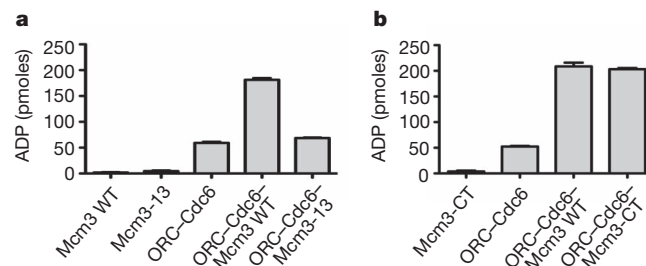
Activation of the ORC-Cdc6 ATPase by the initial recruitment of Mcm3 led us to consider that ATP hydrolysis might play some role early in the Mcm2-7 loading reaction. Supplementary Fig. 9a shows that both ORC and Cdc6 are bound to DNA equally well in ATP and ATP- $\gamma$ S; however, we found that the stable recruitment of Mcm3 we

saw with non-hydrolysable ATP- $\gamma$ S (Supplementary Fig. 3a, f and Fig. 5a, ATP- $\gamma$ S) was lost when hydrolysis was allowed (Fig. 5a, ATP). Similarly, the retention of Mcm3, Mcm5 and Mcm7 seen in ATP- $\gamma$ S in the absence of Cdt1 (for example, Fig. 1) was also lost in the presence of ATP (Fig. 5b, compare lanes 2 and 6). These results indicate that ATP hydrolysis can lead to release of Mcm3 (and associated subunits) when the loading of incomplete Cdt1-Mcm2-7 complexes is attempted.

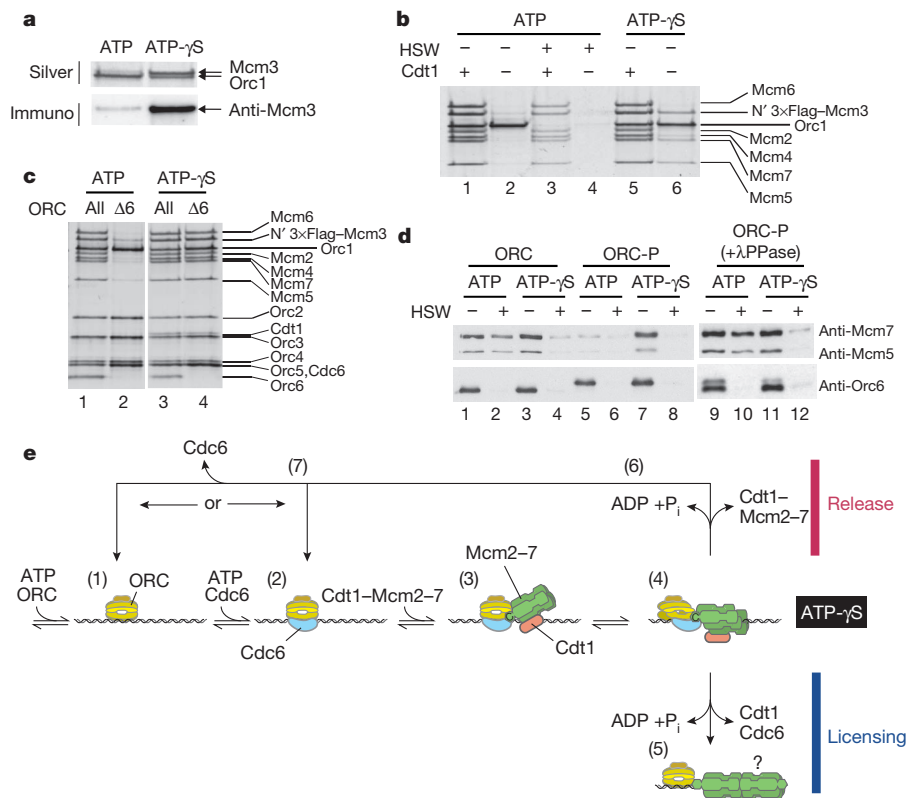
This could be because the remaining Mcm2-7 subunits cannot form stable, topologically closed double hexamers around DNA and are, therefore, released during a futile loading reaction with ATP hydrolysis. ORC is not required for Mcm2-7 to remain bound topologically to the DNA after loading (for example, after a high-salt wash). So, we next looked at the ability of the full, wild-type Cdt1-Mcm2-7 complex to be recruited and loaded by ORC lacking the Orc6 subunit (ORC $\Delta$ 6), which can still bind DNA<sup>14</sup>. ORC $\Delta$ 6 was able to recruit all six Mcm2-7 subunits in the presence of ATP- $\gamma$ S as



**Figure 3 | Both Mcm2-7 hexamers must interact with ORC-Cdc6 through Mcm3.** **a**, Complexes containing wild-type 3xFlag-Mcm3 and MBP-Mcm3-11 were assembled (IN) and mixed in the indicated ratios, with a fixed amount of total Mcm2-7 complex (4 pmol) and tested for loading and recruitment as above. **b**, Loading and recruitment of mixed complexes containing 3xFlag-Mcm3-11 and MBP-Mcm3WT in the indicated ratios.



**Figure 4 | Mcm3 binding activates the ORC-Cdc6 ATPase.** **a**, **b**, The conversion of [ $\alpha$ -<sup>32</sup>P]ATP to [ $\alpha$ -<sup>32</sup>P]ADP was monitored as described in Methods in reactions containing the indicated proteins (Mcm3 wild type; Mcm3-13 mutant; Mcm3-CT, a polypeptide containing the C-terminal 194 amino acid residues of Mcm3; ORC and Cdc6). Error bars depict standard error of the mean from five reactions.



**Figure 5 | ATP hydrolysis by ORC-Cdc6 can promote Mcm2-7 release.**

**a**, Mcm3 was tested for recruitment (low-salt wash) in reactions containing ORC, Cdc6 and either ATP or ATP- $\gamma$ S. Bound proteins were analysed by silver staining (silver) or immunoblotting with anti-Mcm3 antibody (immuno). **b**, Recruitment and loading assays were performed as above with ORC, Cdc6 and Mcm2-7 either with (+) or without (–) Cdt1 in ATP or ATP- $\gamma$ S. Proteins were visualized by silver staining. **c**, Complexes containing either the complete six-subunit ORC complex (All) or a complex lacking Orc6 ( $\Delta$ 6), purified as described in Methods, were tested for recruitment (low-salt wash) in either

ATP or ATP- $\gamma$ S. DNA-bound proteins were analysed by silver staining. **d**, Recruitment and loading assays were performed with unphosphorylated ORC (ORC) or CDK-phosphorylated ORC (ORC-P). In lanes 9–12 ORC-P was previously dephosphorylated with lambda phosphatase ( $\lambda$ PPase). **e**, A revised model for the mechanism of origin licensing which includes an ATPase-dependent release step based on experiments in this manuscript. Details of the model are discussed in the text. The ‘?’ in step (5) refers to the assembly of the second hexamer, which is discussed in the text.

efficiently as the complete ORC complex (Fig. 5c, lanes 3 and 4). Although the recruited complex was slightly more salt-sensitive than the complex recruited with the complete ORC complex (Supplementary Fig. 9b), the only significant difference between Fig. 5c lanes 3 and 4 was the absence of Orc6 in lane 4. Recruitment of Mcm2-7 subunits in the absence of Orc6 still required the C terminus of Mcm3, because it was completely defective in the Mcm3-11 mutant (Supplementary Fig. 9d, lane 4). Notably, although ORC $\Delta$ 6 recruited all six Mcm2-7 subunits in ATP- $\gamma$ S, it was unable to stably recruit any of these subunits in the presence of ATP (compare Fig. 5c, lanes 2 and 4). ORC $\Delta$ 6 and Cdc6 are bound equally well in ATP and ATP- $\gamma$ S (Supplementary Fig. 9a). Thus, although the entire Mcm2-7 ring is efficiently recruited, ATP hydrolysis promotes release of Cdt1-Mcm2-7 in the absence of Orc6.

To rule out the possibility that the non-physiological absence of Orc6 somehow alters the function of ORC-Cdc6, and to examine the role of ATP-dependent Mcm2-7 release *in vivo*, we examined a physiologically relevant situation where ORC has been functionally inactivated. Phosphorylation of Orc6, along with Orc2, by cyclin-dependent kinase (CDK) plays a role in preventing re-replication<sup>15,16</sup> by blocking the ability of ORC to load Mcm2-7<sup>17</sup>. We phosphorylated purified ORC (ORC-P) with the yeast mitotic CDK, Clb2-Cdc28, and examined its ability to recruit and load Mcm2-7. ORC-P cannot load Mcm2-7 (Fig. 5d, compare lanes 2 and 6), but it can recruit Mcm2-7 subunits as well as the unphosphorylated ORC in the presence of ATP- $\gamma$ S (compare lanes 3 and 7). However, ORC-P was unable to accumulate Mcm2-7 subunits in the presence of ATP (compare lanes 1 and 5). Lambda phosphatase treatment of ORC-P restored its ability

to recruit and load Mcm2-7 in ATP (Fig. 5d, lanes 9–12). These results show that, as in the cases of Mcm3 alone, Mcm2-7 without Cdt1, and ORC $\Delta$ 6, the CDK-phosphorylated ORC is competent to recruit Mcm2-7 to ORC-Cdc6, but ATP hydrolysis then promotes Mcm2-7 release.

To provide additional evidence that ATP hydrolysis can promote Mcm2-7 release from ORC $\Delta$ 6, we have looked at Mcm2-7 recruitment in ATP at reduced temperatures, to slow ATP hydrolysis. In contrast to 30 °C (Fig. 5c), the recruitment of Mcm2-7 by ORC $\Delta$ 6 occurred equally well in ATP and ATP- $\gamma$ S at 8 °C (Supplementary Fig. 10). However, when these reactions were shifted to 30 °C after 9 min, recruitment of Mcm2-7 continued in ATP- $\gamma$ S but Mcm2-7 was significantly reduced in the presence of ATP, consistent with the idea that elevated temperature promotes ATP hydrolysis and subsequent Mcm2-7 dissociation.

## Discussion

Our results support a model for origin licensing wherein Cdt1-Mcm2-7 is first recruited to origins by interaction between the C terminus of Mcm3 and ORC-Cdc6 (Fig. 5e step 3). Cdt1 is not required for this initial recruitment, in contrast to previous experiments in crude extracts<sup>17–19</sup>. We suggest, by analogy to other AAA+ proteins<sup>20</sup>, that domain movements in ORC-Cdc6 triggered by Mcm3 binding activate the ATPase of ORC and/or Cdc6 by bringing arginine fingers in contact with ATP binding sites in adjacent subunits (Fig. 5e step 4). We further suggest that these domain movements are transmitted to Mcm2-7, leading to destabilization of Mcm2, Mcm4 and Mcm6 when Cdt1 is missing (Fig. 1A, B). Cdt1 stabilizes Mcm2-7



in this new conformation by a mechanism that does not require its interaction with Orc6 (Fig. 5c). Because both hexamers must interact with ORC–Cdc6 (Fig. 3), our results indicate they are either loaded sequentially or there is more than one Mcm3 binding site in ORC–Cdc6, allowing concerted loading.

When all pre-replicative complex components are present and in the correct post-translational modification state, ATP hydrolysis promotes loading of the ring around DNA (Fig. 5e step 5). However, when criteria for correct licensing are not met, ATP hydrolysis instead is coupled to irreversible dissociation of inappropriate protein assemblies, preventing the accumulation of non-productive complexes (Fig. 5e step 6). Cdc6 is bound to ORCΔ6 in both ATP and ATP-γS (Supplementary Fig. 9e), indicating either that Cdc6 remains bound to ORC after Cdt1–Mcm2–7 release, or that Cdc6 is released, but rebinds quickly (Fig. 5e step 7). This ATPase-dependent quality control mechanism contributes to preventing licensing by CDK-phosphorylated ORC outside of G1 phase, and therefore has a role in ensuring once per cell cycle replication. We note that the block to re-replication in metazoans works largely through inhibition and degradation of Cdt1<sup>21,22</sup>, and our results (Fig. 5b) indicate that this will also prevent licensing via this ATPase-dependent mechanism. Although a great deal is known about the multiple pathways contributing to a robust block to re-assembly of pre-replicative complexes after replication in eukaryotes<sup>23,24</sup>, few quality control mechanisms contributing to efficient pre-replicative complex assembly during G1 phase have been described (see ref. 25). The mechanism we have described here may also have a role during G1 phase to ensure assembly of complete, functional pre-replicative complexes.

## METHODS SUMMARY

All strains are listed in Supplementary Table 1 and strain constructions are described in Methods. Mcm2–7 loading reactions were performed as described (ref. 7 and Methods). After isolation and washing of the beads, DNA was cleaved from beads using irradiation with a 330 nm light source, which does not induce significant levels of DNA crosslinking (see Methods and Supplementary Fig. 11). Proteins were analysed after SDS–PAGE by either silver staining or immunoblotting with a variety of antibodies (Methods). All proteins were expressed in either *S. cerevisiae* or *Escherichia coli* and were purified as described in Methods. Cdt1 and Mcm2–7 expressed in both bacteria and yeast were proficient for loading (Supplementary Fig. 12). ATPase assays were based on the method described in ref. 26 with the modifications described in Methods.

**Full Methods** and any associated references are available in the online version of the paper.

**Received 7 June 2012; accepted 17 January 2013.**

**Published online 10 March 2013.**

- Boos, D., Frigola, J. & Diffley, J. F. X. Activation of the replicative DNA helicase: breaking up is hard to do. *Curr. Opin. Cell Biol.* **24**, 423–430 (2012).
- Tanaka, S. & Araki, H. Regulation of the initiation step of DNA replication by cyclin-dependent kinases. *Chromosoma* **119**, 565–574 (2010).
- Méchal, M. Eukaryotic DNA replication origins: many choices for appropriate answers. *Nature Rev. Mol. Cell Biol.* **11**, 728–738 (2010).
- Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N. & Oda, M. Eukaryotic chromosome DNA replication: where, when, and how? *Annu. Rev. Biochem.* **79**, 89–130 (2010).

- Bell, S. P. & Dutta, A. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* **71**, 333–374 (2002).
- Blow, J. J. & Dutta, A. Preventing re-replication of chromosomal DNA. *Nature Rev. Mol. Cell Biol.* **6**, 476–486 (2005).
- Remus, D. *et al.* Concerted loading of Mcm2–7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**, 719–730 (2009).
- Evrin, C. *et al.* A double-hexameric MCM2–7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl Acad. Sci. USA* **106**, 20240–20245 (2009).
- Gambus, A., Khoudoli, G. A., Jones, R. C. & Blow, J. J. MCM2–7 form double hexamers at licensed origins in *Xenopus* egg extract. *J. Biol. Chem.* **286**, 11855–11864 (2011).
- Speck, C., Chen, Z., Li, H. & Stillman, B. ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nature Struct. Mol. Biol.* **12**, 965–971 (2005).
- Tye, B. K. & Sawyer, S. The hexameric eukaryotic MCM helicase: building symmetry from nonidentical parts. *J. Biol. Chem.* **275**, 34833–34836 (2000).
- Bochman, M. L. & Schwacha, A. The Mcm complex: unwinding the mechanism of a replicative helicase. *Microbiol. Mol. Biol. Rev.* **73**, 652–683 (2009).
- Labib, K., Tercero, J. A. & Diffley, J. F. X. Uninterrupted MCM2–7 function required for DNA replication fork progression. *Science* **288**, 1643–1647 (2000).
- Lee, D. G. & Bell, S. P. Architecture of the yeast origin recognition complex bound to origins of DNA. *Mol. Cell Biol.* **17**, 7159–7168 (1997).
- Nguyen, V. Q., Co, C. & Li, J. J. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* **411**, 1068–1073 (2001).
- Wilmes, G. M. *et al.* Interaction of the S-phase cyclin Clb5 with an ‘RXL’ docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. *Genes Dev.* **18**, 981–991 (2004).
- Chen, S. & Bell, S. P. CDK prevents Mcm2–7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes Dev.* **25**, 363–372 (2011).
- Chen, S., de Vries, M. A. & Bell, S. P. Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2–7 loading. *Genes Dev.* **21**, 2897–2907 (2007).
- Takara, T. J. & Bell, S. P. Multiple Cdt1 molecules act at each origin to load replication-competent Mcm2–7 helicases. *EMBO J.* **30**, 4885–4896 (2011).
- Duderstadt, K. E. & Berger, J. M. AAA+ ATPases in the initiation of DNA replication. *Crit. Rev. Biochem. Mol. Biol.* **43**, 163–187 (2008).
- Fujita, M. Cdt1 revisited: complex and tight regulation during the cell cycle and consequences of deregulation in mammalian cells. *Cell Div.* **1**, 22 (2006).
- Havens, C. G. & Walter, J. C. Mechanism of CRL4<sup>Cdt2</sup>, a PCNA-dependent E3 ubiquitin ligase. *Genes Dev.* **25**, 1568–1582 (2011).
- Diffley, J. F. X. Quality control in the initiation of eukaryotic DNA replication. *Phil. Trans. R. Soc. B* **366**, 3545–3553 (2011).
- Diffley, J. F. X. The many faces of redundancy in DNA replication control. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 135–142 (2010).
- Pasion, S. G. & Forsburg, S. L. Nuclear localization of *Schizosaccharomyces pombe* Mcm2/Cdc19p requires MCM complex assembly. *Mol. Biol. Cell* **10**, 4043–4057 (1999).
- Remus, D., Beall, E. L. & Botchan, M. R. DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC–DNA binding. *EMBO J.* **23**, 897–907 (2004).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We are grateful to N. Cook for help with protein purifications, A. Early and L. Drury for help with strain constructions, G. Coster for help with Mcm3 complementation, S. Mochida and B. Pfander for vectors and K. Labib for antibodies. We also thank members of the Diffley laboratory for critical reading of the manuscript. This work was funded by Cancer Research UK and grants from the Association for International Cancer Research (10-0270) and the European Research Council (249883 – EUKDNAREP).

**Author Contributions** J.F., D.R., A.M. and J.F.X.D. conceived the experiments and wrote the paper. J.F., D.R. and A.M. performed all experiments.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.F.X.D. ([John.Diffley@cancer.org.uk](mailto:John.Diffley@cancer.org.uk)).

## METHODS

**Cdt1–Mcm2–7 overexpressed in yeast (yJF38).** To obtain a yeast strain overexpressing the Cdt1–Mcm2–7 complex, four new vectors were generated using the pRS30<sub>+</sub> series<sup>27</sup>. Making use of the bidirectional inducible *GAL1-10* promoter, the six *MCM* subunits, *CDT1* and *GAL4* genes were cloned and overexpressed in the strain yJF38. *GAL4* is a positive regulator of *GAL* genes in response to galactose, it was therefore overexpressed together with the *MCM2–7* and *CDT1* genes to increase protein yield.

*GAL4* and *MCM2* were amplified using the primer pairs JF1-2 and JF3-4, respectively. *GAL4* was cloned into pRS303 (pJF2.1) and *MCM2* into pRS306 (pJF5.1) between *Sma*I and *Bam*HI sites. *MCM7* and *MCM5* were amplified with JF5-6 and JF7-8 primer pairs and cloned into a *Sma*I site in pRS304 (pJF4.1) and pRS306 (pJF3.1), respectively. The forward primers inserted a unique *Asc*I site at the 5' end of the start codon of these genes. *MCM3*, *MCM6*, *MCM4* and *CDT1* were amplified using the following primer pairs: JF9-10, JF11-12, JF13-14 and JF15-16. The PCR products were cloned between *Spe*I and *Not*I sites into pJF5.1, 4.1, 3.1 and 2.1, respectively, to yield pJF5.2, 4.2, 3.2 and 2.2. In this case, a unique *Sgr*AI site was inserted at the 5' end of their start sites. In addition a 3×Flag tag<sup>28</sup> was inserted using a *Sgr*AI site present upstream of *MCM3* in pJF5.2 (pJF6.2). Finally, the *GAL1-10* promoter was cloned between the *Asc*I and *Sgr*AI sites previously inserted, yielding pJF2, 3, 4 and 6. These vectors were integrated into the strain yJF1 (yJF38).

**Cdt1–Mcm2–7 overexpressed in yeast (yJF35, yJF59 and yJF73).** yJF35 and yJF59 are analogous to yJF38 except that the 3×Flag tag on *MCM3* was inserted at its C terminus in yJF35 and on the C terminus of *CDT1* in yJF59. The vectors integrated into yJF1 were pJF2, 3, 4 and 5. The C-terminal 3×Flag on *MCM3* and *CDT1* was inserted afterwards by PCR-based tagging using pBP83. pJF2, 3, 4 and 39 vectors were integrated into the strain yJF1, yielding yJF73.

**Mcm2–7 overexpressed in yeast (yJF39).** To obtain a Mcm2–7 complex without Cdt1, a strategy similar to the design of yJF38 was followed, except that pJF2 was not integrated into the final strain and the endogenous *CDT1* locus was tandem affinity purification (TAP)-tagged<sup>29</sup> using pJF21, thus producing yJF39.

**ORC overexpressed in yeast (ySD-ORC).** To overexpress ORC in yeast, we modified pJF2, 3 and 5 as follows. In pJF5, *MCM2* and *MCM3* were replaced with calmodulin binding peptide-tobacco etch virus (CBP-TEV)-*ORC1* and -*ORC2*, respectively (pJF19). In pJF2 *GAL4* and *CDT1* were substituted with *ORC3* and *ORC4*, respectively (pJF17). In pJF3, *MCM5* and *MCM4* were replaced with *ORC5* and *ORC6*, respectively (pJF18). All ORC genes were codon optimized. Finally, pJF19, 17 and 18 were integrated into yJF1, producing the strain ySD-ORC.

**ORC(Δ6) overexpressed in yeast (ySD-ORC(Δ6)).** In pJF3, *MCM5* was replaced with *ORC5* and *MCM4* was deleted, producing the vector pJF20. To obtain ORC without Orc6, yJF1 was transformed with pJF17, 19 and pJF20, instead of pJF18. Furthermore, a C-terminal 3×Flag tag was added to the endogenous copy of *ORC6*, using pBP83.

**MCM3 mutants.** *MCM3* mutants were cloned into two different vectors: pMAL-C2P to purify recombinant proteins, and pRS306 for complementation studies in yeast.

pMAL-C2P was a gift from S. Mochida and was derived from pMAL-C2 (NEB) by introducing a PreScission protease site before the *Eco*RI site in the polylinker region<sup>30</sup>. *MCM3* was amplified from *S. cerevisiae* genomic DNA using AM51 and AM52 and cloned into pMAL-C2P using *Xba*I and *Sal*I sites (pAM5). *MCM3* contains a *Kpn*I site at position 2690. Mutant fragments of 231 base pairs in length, between this site and the stop codon were synthesized by GeneArt. The mutated sequences incorporated a *Not*I, followed by a *Sal*I restriction site at the 3' end of the stop codon. Using *Kpn*I and *Sal*I sites, the mutant sequences were cloned into pAM5, to yield pJF27 (Mcm3-11), pJF28 (Mcm3-12), pJF29 (Mcm3-13), pJF32 (Mcm3-16), pJF35 (Mcm3-22) and pJF36 (Mcm3-23).

*MCM2* in pJF5 was removed, to yield pGC003. *MCM3* mutant sequences were excised from pJF27, 28, 29, 32, 35 and 36 vectors using *Avr*II and *Not*I sites, and cloned into pGC003 (pJF8, 9, 10, 13, 15 and 16). *Mcm2* in pJF6 was removed, to yield pJF37. MBP tag was amplified from pMAL-C2P using the primers JF19 and 20 and cloned into pGC003 using *Sgr*AI unique site, originating pJF38. Finally, these vectors were integrated into a *MCM3* degon strain (yKL43<sup>13</sup>), producing the strains yJF63, 64, 65, 66, 69, 71, 72, 78 and 79.

**MCM3 mutants in a diploid background.** Wild-type C-terminal *Mcm3* was amplified from pGC003, C-terminal *Mcm3* mutants were amplified from pJF27 (Mcm3-11), pJF28 (Mcm3-12) and pJF29 (Mcm3-13) using primers JF17 and JF18. The PCR products contained the *URA3* gene and were used to transform the diploid strain W303 (all amplified products were sequenced before transformation). Transformants were selected for on media lacking uracil. These were shown by PCR to be heterozygous for replacement of the C terminus of

*Mcm3* with either wild-type or mutant *Mcm3* cassettes as above. Sporulation and tetrad dissection of these transformants was performed to examine their viability.

**Fragments of MCM3.** A C-terminal fragment of *MCM3* (588 base pairs) was amplified from *S. cerevisiae* genomic DNA using AM54 and AM52. This PCR product was cloned in pMAL-C2P using *Xba*I and *Sal*I sites (pAM6). An N-terminal fragment of *MCM3* (2,331 bp) was amplified from *S. cerevisiae* genomic DNA using AM51 and AM53. This PCR product was cloned in pMAL-C2P using *Xba*I and *Sal*I sites (pAM7).

**Individual MCM subunits expressed in bacteria.** *MCM4* and 5 were amplified by PCR using NC1-2 and NC3-4 primer pairs. These PCR products were cloned between *Nde*I and *Xho*I restriction sites into pET22b, to yield pET22b-*MCM4* and pET22b-*MCM5*. *MCM6* was excised from pET16b-*MCM6* using *Nde*I and *Bam*HI<sup>31</sup> and cloned into pET11a (pET11a-*MCM6*).

**Cdt1 expression in bacteria.** *CDT1* was amplified from *S. cerevisiae* genomic DNA using NC5 and NC6 and cloned into pGEX-6p-1 (GE Healthcare), using *Bam*HI and *Not*I sites (pGEX-6p-1-*CDT1*).

**Cdc6 expression in bacteria.** *S. cerevisiae CDC6* was amplified from pET15b-*CDC6* using the primers AM1 and AM2. The PCR product was cloned between *Bam*HI and *Xho*I restriction sites in pGEX-6p-1 (GE Healthcare), generating pAM3.

**Purification of Cdt1–Mcm2–7 from yJF38, yJF35, yJF59 and yJF73.** 2 l of cells were grown in YP-raffinose at 30 °C to a cell density of  $4 \times 10^7$  cells per ml and arrested for 3 h with 100 ng ml<sup>-1</sup> of alpha-factor. Protein expression was induced by adding galactose (2%) and incubating at 30 °C for 3–4 h. Cells were collected, washed with ice-cold 25 mM HEPES-KOH pH 7.6, 1 M sorbitol, then washed with buffer A (45 mM HEPES-KOH pH 7.6, 0.02% NP-40, 10% glycerol, 5 mM magnesium acetate (Mg(OAc)<sub>2</sub>) and 0.1 M K-glutamate). The pellet was resuspended in 0.5 volumes of buffer A/1 mM DTT/protease inhibitors (Roche) and frozen drop-wise in liquid nitrogen. Frozen drops of cells were crushed using a freezer mill (SPEX CertiPrep 6850 Freezer/Mill) with 6 cycles of 2 min crushing at rate 15. Frozen cell powder was thawed at room temperature, resuspended in 1 volume of buffer A/1 mM DTT/protease inhibitors, and the concentration of K-glutamate was adjusted to 0.5 M. The suspension was centrifuged for 1 h at 50,000 r.p.m. using a Ti70.1 rotor (257,000g). The clear phase was recovered, dialysed twice for 1 h against buffer B (25 mM HEPES-KOH pH 7.6, 0.02% NP-40 and 10% glycerol)/0.1 M K glutamate/5 mM Mg(OAc)<sub>2</sub>, and then centrifuged for 30 min at 50,000 r.p.m. Anti-Flag immunoprecipitation was performed by adding 1 ml packed bead volume of washed anti-Flag M2 agarose (Sigma) to the supernatant and incubating for 1 h at 4 °C. Beads were recovered, and washed with ten bed resin volumes (BVs) of buffer B/0.1 M K glutamate/1 mM ATP/5 mM Mg(OAc)<sub>2</sub>. Elution was performed by adding 0.5 mg ml<sup>-1</sup> 3×Flag peptide. After 20 min at 4 °C, the flow-through was collected, concentrated using a Microcon YM-10, 10000 MWCO (Millipore) and fractionated on a 24 ml Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated in buffer B/0.1 M K acetate. Peak fractions were pooled and aliquoted.

**Purification of Mcm2–7 from yJF39.** yJF39 was purified as described for yJF38. The only difference was a pre-incubation step with IgG Sepharose 6 Fast Flow beads (GE Healthcare), to remove endogenous Cdt1, before the sample was loaded onto the gel filtration column. 100 µl of IgG packed resin was added to the 3×Flag peptide-eluted sample and incubated for 30 min at 4 °C. The unbound fraction was recovered, concentrated and loaded onto a Superdex 200 10/300 GL column (GE Healthcare), as described for yJF39.

**Purification of ORC from ySD-ORC.** 2 l of cells were grown in YP-raffinose at 30 °C to a cell density of  $4 \times 10^7$  cells per ml and arrested for 3 h with 100 ng ml<sup>-1</sup> of alpha-factor. After arrest, protein expression was induced by adding galactose (2%) for 3–4 h at 30 °C. Cells were collected, washed with ice-cold 25 mM HEPES-KOH pH 7.6/1 M sorbitol, then washed with buffer C (25 mM HEPES-KOH pH 7.6, 0.05% NP-40 and 10% glycerol)/0.1 M KCl. The pellet was resuspended in 0.5 volumes of buffer C/0.1 M KCl/1 mM DTT/protease inhibitors (Roche) and frozen drop-wise in liquid nitrogen. Frozen drops of cells were crushed using a freezer mill (SPEX CertiPrep 6850 Freezer/Mill) with 6 cycles of 2 min crushing at rate 15. Frozen cell powder was thawed at room temperature, resuspended in 1 volume of buffer C/0.1 M KCl/1 mM DTT/protease inhibitors, and the concentration of KCl was adjusted to 0.5 M. The suspension was centrifuged for one hour at 50,000 r.p.m. using a Ti70.1 rotor. The clear phase was recovered, and subjected to calmodulin affinity purification by adding 2 mM CaCl<sub>2</sub> and 1.5 ml of packed beads of Calmodulin affinity resin (Stratagene). After 3 h at 4 °C, beads were collected, washed with 10 BVs of buffer C/0.3 M KCl/2 mM CaCl<sub>2</sub>/2 mM β-mercaptoethanol. Elution was performed with buffer C/0.3 M KCl/2 mM EGTA/1 mM EDTA/2 mM β-mercaptoethanol, peak fractions were pooled and subjected to fractionation over a Superdex 200 10/300 GL column (GE Healthcare), pre-equilibrated in buffer C/0.15 mM KCl/2 mM β-mercaptoethanol. Fractions containing ORC were pooled and concentrated



over a 1 ml MonoQ 5/50 GL column (GE Healthcare) using an elution gradient of 0.15–0.5 M KCl over 10 column volumes. Peak fractions containing ORC were pooled, dialysed against buffer C/0.3 M K-acetate and stored in aliquots.

**Purification of ORC( $\Delta 6$ ) from ySD-ORC( $\Delta 6$ ).** To purify ORC without Orc6, a similar protocol to the purification of ORC was followed, except that Flag immunoprecipitation of endogenous Orc6 was included, to remove complexes containing Orc6 (Supplementary Fig. 9c). This was performed as follows: pooled fractions from the gel filtration were incubated for 30 min at 4 °C with 1 ml packed bead volume of anti-Flag M2 agarose (Sigma). Flow-through was collected and concentrated over a MonoQ 5/50 GL column (GE Healthcare), as described for the purification of ORC.

**Proteins expressed in bacteria.** All expression plasmids were transformed into BL21 DE3 Codon+ RIL cells (Stratagene) (unless indicated). 3 l of cells (unless indicated) were grown at 37 °C to a density of  $D_{600\text{ nm}} = 0.5$ –0.8. Cells were chilled on ice, and then IPTG was added to 1 mM. Induction was carried out overnight at 18 °C.

**Purification of MBP–Mcm3.** 0.5 l of cells expressing MBP–Mcm3 were grown as described above. Cells were collected, washed once with ice-cold 25 mM HEPES-KOH pH 7.6/1 M sorbitol, once with buffer D (50 mM Tris-HCl pH 7.5, 0.05% NP-40, 10% glycerol)/1 M NaCl and then the pellet was resuspended in 20 ml of buffer D/1 M NaCl/2 mM  $\beta$ -mercaptoethanol/protease inhibitors (Roche). 50  $\mu$ l of lysozyme (50 mg ml<sup>−1</sup>) was added and the suspension incubated for 20 min at 4 °C. Cells were kept on ice and sonicated 3  $\times$  30 s at 15 microns using a sonicator Soniprep 150 (Sanyo). Lysate was centrifuged for 1 h at 45,000 r.p.m. using a Ti45 rotor (235,400g). The soluble phase was collected and incubated with 2 ml packed amylose bead volume (NEB) at 4 °C for 1 h. Beads were washed with ten BVs of buffer D/0.3 M NaCl/2 mM  $\beta$ -mercaptoethanol. Elution was performed with buffer D/0.3 M NaCl/2 mM  $\beta$ -mercaptoethanol/10 mM maltose. Peak fractions were pooled and 0.5 ml were subjected to fractionation over a Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated in buffer D/0.3 M NaCl/2 mM  $\beta$ -mercaptoethanol. The peak fraction, of correct molecular weight, was collected and subjected to PreScission protease cleavage with 100 units of PreScission protease (GE healthcare). After 30 min at room temperature, 1 volume of buffer D/2 mM  $\beta$ -mercaptoethanol was added, and the sample was fractionated over a 0.15 ml MonoQ PC1.6/5 column (GE Healthcare), using an elution gradient of 0.15–0.7 M NaCl over 50 column volumes. Peak fractions containing untagged Mcm3 proteins were kept. The same protocol was used to purify Mcm3 fragments and mutants.

**Purification of Mcm2, Mcm7 and Mcm3.** To purify these proteins, we followed the protocol described in ref. 31. The starting cultures were 3 l in volume.

**Purification of Mcm5.** Cells were grown and expressed as described above. Cells were collected, washed once with ice-cold 25 mM HEPES-KOH pH 7.6/1 M sorbitol, once with buffer E (50 mM Tris-HCl pH 7.5, 0.05% NP-40, 10% glycerol, 1 mM EDTA and 1 mM DTT)/1 M NaCl, and the pellet was resuspended in 40 ml of buffer E/1 M NaCl/protease inhibitors (Roche). Cells were sonicated and the lysate centrifuged as described for the purification of MBP–Mcm3. The soluble fraction from centrifugation was treated with 0.3 g ml<sup>−1</sup> ammonium sulphate, and stirred for 20 min at 4 °C. The mixture was centrifuged at 17,000 r.p.m. (34,600g) for 20 min. The pellet was resuspended in 40 ml of buffer E containing 0.25 g ml<sup>−1</sup> of ammonium sulphate and centrifuged as before. This was repeated with 0.20 mg ml<sup>−1</sup> ammonium sulphate. The pellet was resuspended in 30 ml of buffer E and dialysed against buffer E for one hour at 4 °C. After checking that the conductivity of the sample was below that of buffer E/0.1 M NaCl, the sample was applied to a 5 ml FF Q column (GE Healthcare), pre-equilibrated in buffer E/0.1 M NaCl. Elution was done using a gradient of 0.1 M to 0.5 M NaCl over 10 column volumes. Peak fractions were pooled and diluted down to a conductivity below that of buffer E/0.1 M NaCl. The sample was then applied to a 5 ml Heparin column (GE healthcare) pre-equilibrated in buffer E/0.1 M NaCl. Flow through was collected and subjected to fractionation over an 8 ml MonoQ column (GE Healthcare). Elution was performed using a gradient of 0.1 M to 0.5 M NaCl over 20 column volumes. Peak fractions were collected and subjected to fractionation over a Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated with buffer E/0.3 M NaCl. Peak fractions were pooled, concentrated using a Microcon YM-10, 10000 MWCO (Millipore) and aliquoted.

**Purification of Mcm4.** 3 l of expressing cells were grown and we followed the purification process described in ref. 31. At the end of the described protocol, we introduced an extra gel filtration step using a Superdex 200 10/300 GL column (GE Healthcare), pre-equilibrated in buffer E/0.3 M NaCl.

**Purification of Mcm6.** To purify Mcm6, the same protocol as that described for Mcm5 was followed, the difference being that Mcm6 was bound to a 5 ml heparin column (GE Healthcare). Bound protein was eluted using a gradient of 0.1 to 0.5 M NaCl over 20 column volumes. Peak fractions were pooled and the conductivity was reduced to below that of buffer E/0.1 M NaCl. The sample

was then applied to an 8 ml MonoQ column (GE Healthcare), as described in the purification of Mcm5.

**Purification of Cdt1.** Expressing cells were grown and induced as described before. Cells were collected, washed with PBS pH 7.5, and then washed with buffer F (PBS pH 7.5, 10% glycerol, 1 mM DTT and 0.3 M NaCl). The pellet was resuspended in 40 ml of buffer F/protease inhibitors (Roche) and sonicated 3  $\times$  30 s using a sonicator Soniprep 150 (Sanyo) at 15 microns. Lysate was centrifuged at 45,000 r.p.m. (235,400g) for 60 min using a Ti45 rotor. 2 ml of packed bead volume glutathione Sepharose 4Fast Flow (GE Healthcare) were added to the soluble phase and incubated for 2.5 h at 4 °C. The solution was centrifuged at 3,000 r.p.m. (1,400g) for 3 min, beads were collected, 10 column volumes of buffer F were added, and the mixture was incubated for 10 min at 4 °C. This wash step was repeated with 20 column volumes of buffer F, and 10 column volumes of buffer E/0.3 M NaCl. The pellet was then resuspended in 2 ml of buffer E/0.3 M NaCl. 125 units of PreScission Protease (GE Healthcare) were added and the mixture was incubated overnight at 4 °C. The flow through was collected and concentrated using a Microcon YM-10, 10000 MWCO (Millipore), then the sample was loaded onto a gel filtration column (Superdex 200 10/300 GL column (GE Healthcare)) pre-equilibrated in buffer B/1 mM EDTA/1 mM DTT/0.3 M NaCl. Peak fractions were pooled, concentrated and aliquoted.

**Formation of MCM complexes from individually purified subunits.** MCM complexes were formed by combining 10  $\mu$ g of individually purified subunits and fractionated over a Superdex 200 PC 3.2/30 column (GE Healthcare), pre-equilibrated in buffer B/0.1 M K acetate. High molecular weight fractions containing Cdt1–Mcm2–7 complex were pooled, concentrated and aliquoted. Supplementary Fig. 12 shows that this complex can be loaded as efficiently as the complex purified from yeast.

**Purification of GST–Cdc6.** This protocol is modified from that described in ref. 10. 1 l of expressing cells were grown at 37 °C to  $D_{600\text{ nm}} = 0.6$ , then induced with 0.5 mM IPTG for 5 h at 18 °C. Cells were collected at 6,000 r.p.m. (6,100g) in an SLA-3000 rotor (Sorvall) for 10 min. Pellets were resuspended in 50 ml buffer G (50 mM K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub>, pH 7.5, 5 mM MgCl<sub>2</sub>, 1% Triton X-100 and 1 mM DTT)/2 mM ATP/0.15 M KOAc/protease inhibitors (Roche) and 100  $\mu$ g ml<sup>−1</sup> lysozyme added. The mixture was incubated at 4 °C for 30 min and sonicated for 2 min (5 s off, 5 s on) at 15 microns. The suspension was centrifuged at 15,000 r.p.m. (27,000g) for 15 min in a SS34 rotor (Sorvall) and the supernatant transferred to 2 ml bed resin glutathione Sepharose (GE Healthcare). This was rotated at 4 °C for 3 h. Glutathione beads and bound proteins were recovered and washed with 20 column volumes of buffer G/0.15 M KOAc/2 mM ATP. A 50% slurry with buffer G/0.15 M KOAc/2 mM ATP was made and 100 units preScission protease (GE Healthcare) added. The mixture was incubated for 2 h at 4 °C. The flow-through was recovered and the concentration of KOAc diluted to 75 mM with buffer G/2 mM ATP. This was incubated with 2 ml bed resin hydroxyapatite prewashed in buffer G/0.075 M KOAc/2 mM ATP. The protein–hydroxyapatite was washed with 5 BVs of buffer G/2 mM ATP and then washed with 5 BVs of buffer G/0.15 M KOAc/15% glycerol. Cdc6 was eluted with buffer G/0.4 M KOAc/15% glycerol. Peak fractions were pooled and concentrated using a Centricon Plus-20 Centrifugal Filter (Millipore), then aliquoted.

**Purification of ORC-P.** To purify CDK-phosphorylated ORC (ORC-P), ORC was isolated from 50 l of alpha-factor-arrested G1 yeast (YDR11) cells via calmodulin pull-down as described<sup>8</sup>, incubated with TEV protease overnight at 4 °C to remove the tag, dialysed against 25 mM HEPES-KOH pH 7.6/0.1 M KCl/5 mM Mg(OAc)<sub>2</sub>/0.02% NP40/10% glycerol/1 mM DTT, and concentrated to a final volume of 2 ml.

In a parallel preparation, 50 l of YDR12 (*MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100 cdc15-2 bar1::kanMX pep4::HIS3 ura3::Pgall,10-CDC6-TAPtcp (URA3)*) were grown at 25 °C in YP- raffinose to 2  $\times$  10<sup>7</sup> cells per ml. Cdc6–TAP<sup>TCP</sup> expression was induced by addition of 2% galactose for 8 h at 25 °C, upon which the cells arrest with a long-budded phenotype and with replicated DNA in G2/M phase. The cells were collected by centrifugation, washed twice with cold buffer containing 25 mM HEPES-KOH pH 7.6/1 M sorbitol, once with cold buffer containing 45 mM HEPES-KOH pH 7.6/0.1 M KCl/0.02% NP-40/10% glycerol, resuspended in 0.5 volume of packed cell volume of buffer containing 45 mM HEPES-KOH pH 7.6/0.1 M KCl/0.02% NP-40/10% glycerol/2 mM DTT/2 $\times$  complete protease inhibitor cocktail (Roche), and the resulting cell suspension frozen drop-wise directly in liquid nitrogen. The frozen cell suspension was crushed using a SPEX 6870 Freezer/Mill, the resulting frozen powder thawed on ice, diluted with 1 volume of 45 mM HEPES-KOH pH 7.6/0.1 M KCl/0.02% NP-40/10% glycerol/1 mM DTT, the salt concentration adjusted to 0.3 M KCl, and the resulting lysate clarified by centrifugation in a Beckman 45 Ti rotor at 42,000 r.p.m. (205,000g) for 1 h. The clarified extract was supplemented with 2 mM CaCl<sub>2</sub>, and a complex containing Cdc6–TAP<sup>TCP</sup>–Cib2–Cdc28–Cks1 was isolated from this extract using calmodulin affinity

purification in buffer containing 45 mM HEPES-KOH pH 7.6/0.3 M KCl/0.02% NP-40/10% glycerol/1 mM DTT. This partially purified complex was incubated with TEV protease overnight at 4 °C to remove the tag from Cdc6, dialysed against 25 mM HEPES-KOH pH 7.6/0.1 M KCl/5 mM Mg(OAc)<sub>2</sub>/0.02% NP40/10% glycerol/1 mM DTT, and concentrated to a final volume of 2 ml.

The concentrated calmodulin-purified and TEV-protease-digested ORC and Cdc6-Clb2-Cdc28-Cks1 fractions were combined, supplemented with 3 mM ATP, and incubated for 45 min at 30 °C. The phosphorylated ORC (ORC-P) resulting from this reaction was re-isolated from the reaction by gel-filtration on a 120 ml Superdex 200 column in a buffer containing 25 mM HEPES-KOH pH 7.6/0.1 M KCl/0.02% NP40/10% glycerol/1 mM DTT, followed by fractionation on a 1 ml MonoQ ion-exchange column as described<sup>8</sup>.

**Loading reaction.** In this study, four main differences to the conditions described in ref. 7 were introduced in the loading reactions. K-glutamate was substituted with K-acetate in the binding and washing buffers. For silver staining purposes we increased the amount of protein used in the assays by four fold compared to ref. 7. In addition, ARS305 was amplified using an oligonucleotide primer with a photocleavable biotin (Integrated DNA technologies) at one end, as described in ref. 32. We have optimised the photocleavage to minimise DNA damage by irradiating for 10 min at 330 nm (See Supplementary Fig. 11). And finally, 2.5 pmol of DNA molecules have been used, instead of the 1 pmol used before.

**Loading assays with ORC and ORC-P.** Mcm2–7 loading onto immobilized linear 1 kb ARS305-containing DNA using 50 nM purified ORC or ORC-P was performed as described<sup>8</sup>. To test the effect of ORC-P dephosphorylation by lambda-phosphatase, 20 pmol of purified ORC-P at a concentration of 0.67 mM was supplemented with 20 mM MnCl<sub>2</sub>, and incubated for 20 min at 30 °C with either 400 units of l-phosphatase (NEB) or a buffer control before addition of the treated ORC-P to the loading reaction.

**ATPase assays.** Based on the method described in ref. 26 with the following modifications. Reactions were carried out at 30 °C for 20 min, buffer contained 25 mM HEPES-KOH, pH 7.6/0.1% NP-40/5 mM Mg(OAc)<sub>2</sub>/1 mM EDTA/1 mM EGTA/100 mM K-acetate/5% glycerol/1 mM DTT/100 μM ATP (including 2.5 μCi of [α-<sup>32</sup>P]ATP). Each reaction contained 2.5 pmol of 1 kb linear

ARS305 DNA. Where indicated, 2.5 pmol of each protein were also included in these reactions. Reactions were stopped by spotting 1 μl of each reaction on PEI-cellulose TLC plates (CamLab). The cellulose membrane was developed in 0.6 M Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> pH 3.5, and quantified on a Phosphorimager (GE Healthcare).

**Antibodies for western blot analysis.** Anti-Mcm2 (yN-19, sc-6680, Santa Cruz), anti-Mcm7 (yN-19, sc-6688, Santa Cruz), anti-Mcm4 (yC-19, sc-6685, Santa Cruz), anti-Mcm5 (yC-19, sc-6687, Santa Cruz), anti-Orc6 (SB49), anti-Cdc6 (98H/5), anti-Cdt1<sup>33</sup>, anti-Flag M2-peroxidase (Sigma), anti-PAP (peroxidase-anti-peroxidase)-HRP (horseradish peroxidase) (Sigma), anti-MBP (maltose binding protein)-HRP (NEB). Antibodies against Mcm6 and Mcm3 were gifts from the Labib laboratory<sup>34</sup>.

27. Sikorski, R. S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19–27 (1989).
28. Gelbart, M. E., Rechsteiner, T., Richmond, T. J. & Tsukiyama, T. Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. *Mol. Cell. Biol.* **21**, 2098–2106 (2001).
29. Puig, O. *et al.* The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**, 218–229 (2001).
30. Mochida, S., Ikeo, S., Gannon, J. & Hunt, T. Regulated activity of PP2A-B55δ is crucial for controlling entry into and exit from mitosis in *Xenopus* egg extracts. *EMBO J.* **28**, 2777–2785 (2009).
31. Davey, M. J., Indiani, C. & O'Donnell, M. Reconstitution of the Mcm2–7p heterohexamers, subunit arrangement, and ATP site architecture. *J. Biol. Chem.* **278**, 4491–4499 (2003).
32. Tsakraklides, V. & Bell, S. P. Dynamics of pre-replicative complex assembly. *J. Biol. Chem.* **285**, 9437–9443 (2010).
33. Tanaka, S. & Diffley, J. F. X. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2–7 during G1 phase. *Nature Cell Biol.* **4**, 198–207 (2002).
34. Gambus, A. *et al.* GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nature Cell Biol.* **8**, 358–366 (2006).



# Dusty starburst galaxies in the early Universe as revealed by gravitational lensing

J. D. Vieira<sup>1</sup>, D. P. Marrone<sup>2</sup>, S. C. Chapman<sup>3,4</sup>, C. De Breuck<sup>5</sup>, Y. D. Hezaveh<sup>6</sup>, A. Weiß<sup>7</sup>, J. E. Aguirre<sup>8</sup>, K. A. Aird<sup>9</sup>, M. Aravena<sup>5</sup>, M. L. N. Ashby<sup>10</sup>, M. Bayliss<sup>11</sup>, B. A. Benson<sup>12,13</sup>, A. D. Biggs<sup>5</sup>, L. E. Bleem<sup>12,14</sup>, J. J. Bock<sup>1,15</sup>, M. Bothwell<sup>2</sup>, C. M. Bradford<sup>15</sup>, M. Brodwin<sup>16</sup>, J. E. Carlstrom<sup>12,13,14,17,18</sup>, C. L. Chang<sup>12,13,18</sup>, T. M. Crawford<sup>12,17</sup>, A. T. Crites<sup>12,17</sup>, T. de Haan<sup>6</sup>, M. A. Dobbs<sup>6</sup>, E. B. Fomalont<sup>19</sup>, C. D. Fassnacht<sup>20</sup>, E. M. George<sup>21</sup>, M. D. Gladders<sup>12,17</sup>, A. H. Gonzalez<sup>22</sup>, T. R. Greve<sup>23</sup>, B. Gullberg<sup>5</sup>, N. W. Halverson<sup>24</sup>, F. W. High<sup>12,17</sup>, G. P. Holder<sup>6</sup>, W. L. Holzapfel<sup>21</sup>, S. Hoover<sup>12,13</sup>, J. D. Hrubes<sup>9</sup>, T. R. Hunter<sup>19</sup>, R. Keisler<sup>12,14</sup>, A. T. Lee<sup>21,25</sup>, E. M. Leitch<sup>12,17</sup>, M. Lueker<sup>1</sup>, D. Luong-Van<sup>9</sup>, M. Malkan<sup>26</sup>, V. McIntyre<sup>27</sup>, J. J. McMahon<sup>12,13,28</sup>, J. Mehl<sup>12,17</sup>, K. M. Menten<sup>7</sup>, S. S. Meyer<sup>12,13,14,17</sup>, L. M. Mocanu<sup>12,17</sup>, E. J. Murphy<sup>29</sup>, T. Natoli<sup>12,14</sup>, S. Padin<sup>1,12,17</sup>, T. Plagge<sup>12,17</sup>, C. L. Reichardt<sup>21</sup>, A. Rest<sup>30</sup>, J. Ruel<sup>11</sup>, J. E. Ruhl<sup>31</sup>, K. Sharon<sup>12,17,32</sup>, K. K. Schaffer<sup>12,33</sup>, L. Shaw<sup>6,34</sup>, E. Shirokoff<sup>1</sup>, J. S. Spilker<sup>2</sup>, B. Stalder<sup>10</sup>, Z. Staniszewski<sup>1,31</sup>, A. A. Stark<sup>10</sup>, K. Story<sup>12,14</sup>, K. Vanderlinde<sup>6</sup>, N. Welikala<sup>35</sup> & R. Williamson<sup>12,17</sup>

**In the past decade, our understanding of galaxy evolution has been revolutionized by the discovery that luminous, dusty starburst galaxies were 1,000 times more abundant in the early Universe than at present<sup>1,2</sup>. It has, however, been difficult to measure the complete redshift distribution of these objects, especially at the highest redshifts ( $z > 4$ ). Here we report a redshift survey at a wavelength of three millimetres, targeting carbon monoxide line emission from the star-forming molecular gas in the direction of extraordinarily bright millimetre-wave-selected sources. High-resolution imaging demonstrates that these sources are strongly gravitationally lensed by foreground galaxies. We detect spectral lines in 23 out of 26 sources and multiple lines in 12 of those 23 sources, from which we obtain robust, unambiguous redshifts. At least 10 of the sources are found to lie at  $z > 4$ , indicating that the fraction of dusty starburst galaxies at high redshifts is greater than previously thought. Models of lens geometries in the sample indicate that the background objects are ultra-luminous infrared galaxies, powered by extreme bursts of star formation.**

We constructed a catalogue of high-redshift ( $z > 1$ ) galaxy candidates from the first 1,300 square degrees of the South Pole Telescope (SPT)<sup>3</sup> survey by selecting sources with dust-like spectral indexes in the 1.4 and 2.0 mm SPT bands<sup>4</sup>. A remarkable aspect of selecting sources based on their flux at millimetre wavelengths is the so-called negative  $k$ -correction<sup>5</sup>, whereby cosmological dimming is compensated by the steeply rising dust spectrum as the source redshift increases. As a result, a millimetre-wave-selected sample should draw from the redshift distribution of dusty starburst galaxies with little bias over the entire redshift range in which they are expected to exist. To isolate the high-redshift, dusty-spectrum galaxy population, sources with

counterparts in the IRAS Faint Source Catalog<sup>6</sup> (typically  $z < 0.03$ ) were removed, and those with counterparts in the 843 MHz Sydney University Molonglo Sky Survey<sup>7</sup> were removed to exclude sources with strong synchrotron emission (for example, flat-spectrum radio quasars) that may have passed the spectral index cut. A sample of 47 sources with 1.4-mm flux density  $> 20$  mJy and accurate positions were selected for high-resolution imaging with the Atacama Large Millimeter/submillimeter Array (ALMA). Our ALMA spectroscopic observations targeted a sample of 26 sources, all but two of which are in the imaging sample (see Supplementary Information). These objects are among the brightest dusty-spectrum sources in the  $z > 0.1$  extragalactic sky at millimetre wavelengths.

Gravitationally lensed sources are expected to predominate in samples of the very brightest dusty galaxies because of the rarity of unlensed dusty starburst galaxies at these flux levels<sup>8–10</sup>. Massive elliptical galaxies, acting as lenses, will have Einstein radii as large as  $2''$  and may magnify background galaxies by factors of 10 or more. To confirm the lensing hypothesis and determine magnifications, we imaged 47 SPT sources with ALMA at 870  $\mu$ m in two array configurations, which provide angular resolutions of  $1.5''$  and  $0.5''$  (full-width at half-maximum). A sample of these objects with infrared imaging, spectroscopic redshifts and resolved structure is shown in Fig. 1. Integration times of only one minute per source are adequate to show that most sources are resolved into arcs or Einstein rings—hallmarks of gravitational lensing. For all sources for which we have infrared and submillimetre imaging, as well as spectroscopic redshifts, the emission detected by ALMA coincides with massive foreground galaxies or galaxy groups/clusters, but is spatially distinct and at drastically different redshifts (see Fig. 2 and Supplementary Fig. 3). Using a modelling

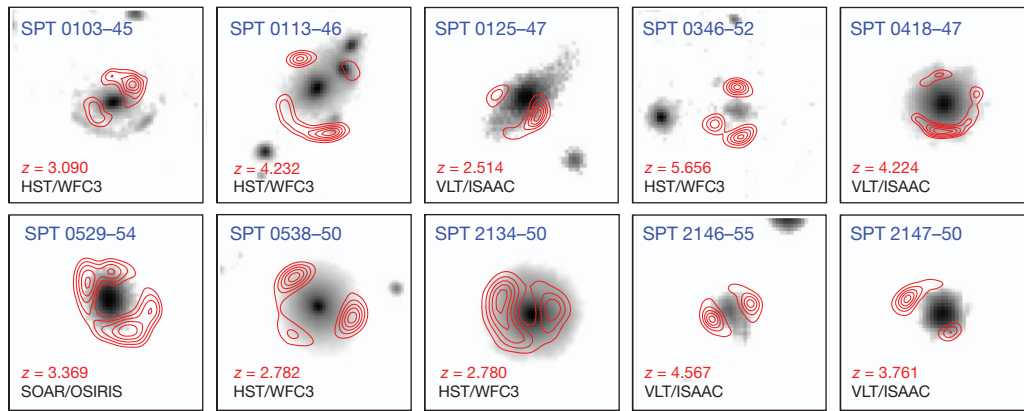
<sup>1</sup>California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA. <sup>2</sup>Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, Arizona 85721, USA.

<sup>3</sup>Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia B3H 3J5, Canada. <sup>4</sup>Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK.

<sup>5</sup>European Southern Observatory, Karl-Schwarzschild Strasse, D-85748 Garching bei München, Germany. <sup>6</sup>Department of Physics, McGill University, 3600 Rue University, Montreal, Quebec H3A 2T8, Canada. <sup>7</sup>Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany. <sup>8</sup>University of Pennsylvania, 209 South 33rd Street, Philadelphia, Pennsylvania 19104, USA. <sup>9</sup>University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA. <sup>10</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA. <sup>11</sup>Department of Physics, Harvard University, 17 Oxford Street, Cambridge, Massachusetts 02138, USA. <sup>12</sup>Kavli Institute for Cosmological Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA.

<sup>13</sup>Enrico Fermi Institute, University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA. <sup>14</sup>Department of Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA.

<sup>15</sup>Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, California 91109, USA. <sup>16</sup>Department of Physics and Astronomy, University of Missouri, 5110 Rockhill Road, Kansas City, Missouri 64110, USA. <sup>17</sup>Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA. <sup>18</sup>Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, USA. <sup>19</sup>National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, Virginia 22903, USA. <sup>20</sup>Department of Physics, University of California, One Shields Avenue, Davis, California 95616, USA. <sup>21</sup>Department of Physics, University of California, Berkeley, California 94720, USA. <sup>22</sup>Department of Astronomy, University of Florida, Gainesville, Florida 32611, USA. <sup>23</sup>Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK. <sup>24</sup>Department of Astrophysical and Planetary Sciences and Department of Physics, University of Colorado, Boulder, Colorado 80309, USA. <sup>25</sup>Physics Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. <sup>26</sup>Department of Physics and Astronomy, University of California, Los Angeles, California 90095-1547, USA. <sup>27</sup>Australia Telescope National Facility, CSIRO, Epping, New South Wales 1710, Australia. <sup>28</sup>Department of Physics, University of Michigan, 450 Church Street, Ann Arbor, Michigan 48109, USA. <sup>29</sup>Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, California 91101, USA. <sup>30</sup>Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. <sup>31</sup>Physics Department, Center for Education and Research in Cosmology and Astrophysics, Case Western Reserve University, Cleveland, Ohio 44106, USA. <sup>32</sup>Department of Astronomy, University of Michigan, 500 Church Street, Ann Arbor, Michigan 48109, USA. <sup>33</sup>Liberal Arts Department, School of the Art Institute of Chicago, 112 South Michigan Avenue, Chicago, Illinois 60603, USA. <sup>34</sup>Department of Physics, Yale University, PO Box 208210, New Haven, Connecticut 06520-8120, USA. <sup>35</sup>Institut d'Astrophysique Spatiale, Bâtiment 121, Université Paris-Sud XI et CNRS, 91405 Orsay Cedex, France.



**Figure 1 | Near-infrared and ALMA submillimetre-wavelength images of SPT targets.** Images are  $8'' \times 8''$ . We show 10 sources for which we have confirmed ALMA spectroscopic redshifts, deep near-infrared (NIR) imaging, and well-resolved structure in the ALMA 870  $\mu\text{m}$  imaging; source names are in blue in each panel. The greyscale images are NIR exposures from the Hubble Space Telescope Wide Field Camera 3 (HST/WFC3, co-added F160W and F110W filters), the Very Large Telescope Infrared Spectrometer and Array Camera (VLT/ISAAC:  $K_s$  band) or the Southern Astrophysical Research Telescope Ohio State Infrared Imager/Spectrometer (SOAR/OSIRIS:  $K_s$  band), and trace the starlight from the foreground lensing galaxy. The NIR images are shown with logarithmic stretch, and each panel shows at bottom left in black the telescope/instrument used to obtain the image. The red contours are ALMA

870  $\mu\text{m}$  imaging showing the background source structure, clearly indicative of strong lensing from galaxy-scale haloes. In all cases, the contours start at  $5\sigma$  and are equally spaced up to 90% of the peak significance, which ranges from 12 to 35. Spectroscopic redshifts of the background sources are shown in red in each panel, above the NIR telescope/instrument names. The ALMA exposures were approximately 2-min integrations, roughly equally divided between the compact and extended array configurations. The resulting resolution is  $0.5''$ . SPT 0103-45 shows a rare lensing configuration of one lens and two background sources at different redshifts, one visible with ALMA and one with HST. SPT 0346-52, with a CO-derived redshift of  $z = 5.656$ , is among the highest-redshift starbursts known. (See Supplementary Information for more details.)

procedure that treats the interferometer data in their native measurement space, rather than through reconstructed sky images, to simultaneously determine the source/lens configuration and correct for antenna-based phase errors<sup>11</sup>, we are able to determine magnifications and derive intrinsic luminosities for our sources. Complete models of four lenses<sup>11</sup>, as well as preliminary models of eight more, indicate lensing magnifications between 4 and 22. After correcting for the magnification, these sources are extremely luminous—more than  $10^{12}$  times solar luminosity ( $L_\odot$ ) and sometimes  $>10^{13}L_\odot$ —implying star-formation rates in excess of  $500 M_\odot \text{ yr}^{-1}$ .

Obtaining spectroscopic redshifts for high-redshift, dusty starburst galaxies has been notoriously difficult. To date, most spectroscopic redshift measurements have come from the rest-frame ultraviolet and optical wavebands after multi-wavelength counterpart identification<sup>2,12,13</sup>. These observations are difficult, owing to the extinction of the ultraviolet light by the dust itself, the cosmological dimming, and the ambiguity in the association of the dust emission with multiple sources of optical emission visible in deep observations. A much more direct method to determine redshifts of starburst galaxies, particularly at high redshift, is through observations of molecular emission associated with their dusty star-forming regions. The millimetre and submillimetre transitions of molecular carbon monoxide (CO) and neutral carbon (C I) are well-suited for this purpose<sup>14</sup>. These emission lines are a major source of cooling for the warm molecular gas fuelling the star formation, and can thus be related unambiguously to the submillimetre continuum source<sup>15</sup>. Until recently, bandwidth and sensitivity limitations made this approach time-intensive. The combination of ALMA—even with its restricted early science capabilities and only 16 antennas—and a unique sample of extraordinarily bright millimetre sources has changed this situation greatly, allowing us to undertake a sensitive, systematic search for molecular and atomic lines across broad swaths of redshift space at  $z > 1$ .

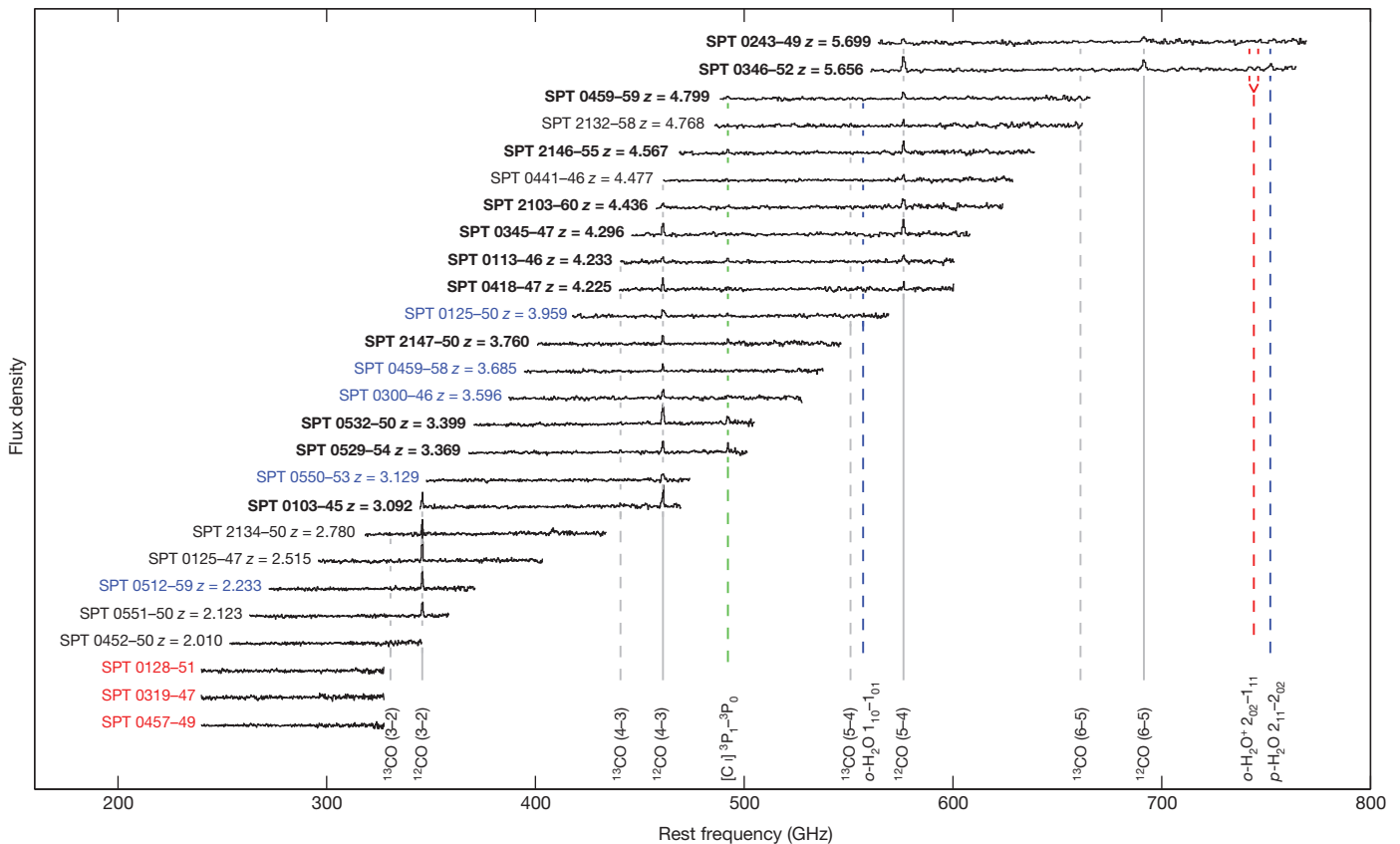
We conducted a redshift search in the 3 mm atmospheric transmission window with ALMA using five spectral tunings to cover 84.2–114.9 GHz. For  $z > 1$ , at least one CO line will fall in this frequency range, except for a small redshift ‘desert’ ( $1.74 < z < 2.00$ ). For sources at  $z > 3$ , multiple transitions (such as rotational transitions of CO and  $\text{C I}(^3\text{P}_1 \rightarrow ^3\text{P}_0)$ ) are redshifted into the observing band, allowing for an unambiguous redshift determination. We find one or more spectral

features in 23 of 26 SPT-selected sources. The detections comprise 44 emission line features, which we identify as redshifted emission from molecular transitions of  $^{12}\text{CO}$ ,  $^{13}\text{CO}$ ,  $\text{H}_2\text{O}$  and  $\text{H}_2\text{O}^+$ , and a C I fine structure line. The spectra of all sources are shown in Fig. 2. For 18 of the sources we are able to infer unique redshift solutions, either from ALMA data alone (12), or with the addition of data from the Very Large Telescope and/or the Atacama Pathfinder Experiment telescope (6). With the 10  $z > 4$  objects discovered here, we have more than doubled the number of spectroscopically confirmed, ultra-luminous galaxies discovered at  $z > 4$  in millimetre/submillimetre surveys in the literature (of which just nine have been reported previously<sup>13,14,16–21</sup>). Two sources are at  $z = 5.7$ , placing them among the most distant ultra-luminous starburst galaxies known.

The SPT dusty galaxy redshift sample comprises 28 sources, as we include an additional two SPT sources with spectroscopic redshifts<sup>22</sup> that would have been included in the ALMA program had their redshifts not already been determined. Of the 26 ALMA targets, three lack a spectral line feature in the ALMA band. We tentatively and conservatively place these at  $z = 1.85$ , in the middle of the  $z = 1.74$ – $2.00$  redshift desert, though it is also possible that they are located at very high redshift or have anomalously faint CO lines. For the five sources for which only a single emission line is found, only two or three redshifts are possible (corresponding to two choices of CO transition) after excluding redshift choices for which the implied dust temperature—derived from our extensive millimetre/submillimetre photometric coverage (provided by ALMA 3 mm, SPT 2 and 1 mm, APEX/LABOCA 870  $\mu\text{m}$  and Herschel/SPIRE 500, 350, 250  $\mu\text{m}$  observations<sup>23</sup>)—is inconsistent with the range seen in other luminous galaxies<sup>22</sup>. For these sources, we adopt the redshift corresponding to the dust temperature closest to the median dust temperature in the unambiguous spectroscopic sample, as shown in Fig. 2.

The cumulative distribution function of all redshifts in this sample is shown in Fig. 3. The median redshift of our full sample is  $z_{\text{med}} = 3.6$ . The redshift distribution of SPT sources with millimetre spectroscopic redshifts is in sharp contrast to that of radio-identified starbursts with optical spectroscopic redshifts, which have a significantly lower median redshift of  $z_{\text{med}} = 2.2$ , and for which only 15–20% of the population is expected to be at  $z > 3$  (ref. 2). Part of this difference can be attributed to the high flux threshold of the original SPT



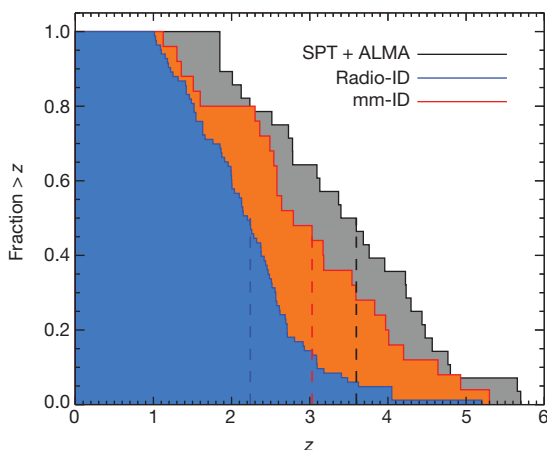


**Figure 2 | ALMA 3 mm spectra of 26 SPT sources.** The vertical axis is observed flux density in units of mJy, with 30 mJy offsets between sources for clarity. Spectra are continuum-subtracted. The strong CO lines are indicative of dust-enshrouded active star formation. The spectra are labelled by source and redshift. Black labels indicate unambiguous redshifts (18), with the subset in bold font (12) having been derived from the ALMA data alone. Sources labelled in blue (5) are plotted at the most likely redshift of multiple options, based on the dust temperature derived from extensive far-infrared photometry. Three sources with no lines detected are placed at  $z = 1.85$ , in the middle of the

redshift range for which we expect no strong lines, and labelled in red. Total integration times for each source were roughly ten minutes. The synthesized beam size ranges from  $7'' \times 5''$  to  $5'' \times 3''$  over the frequency range of the search, which is inadequate to spatially resolve the velocity structure of the lensed sources. Transitions of species detected in at least one source are indicated by vertical lines. Rotational transitions of the  $^{12}\text{CO}$  (solid) and  $^{13}\text{CO}$  (dashed) isotopologues are shown in grey. Water lines are marked by blue dashed lines, ionized water lines by red dashed lines, and atomic carbon ([C I]) by the green dashed line.

selection, which effectively requires that the sources be gravitationally lensed. A much smaller total volume is lensed at  $z < 1$  than at higher redshift, and, as expected, we do not find any such sources in the SPT sample<sup>23</sup>. However, if we only compare sources at  $z > 2$  (the lowest confirmed spectroscopic redshift in the SPT sample), the median redshift of the radio-identified sample is still significantly lower (2.6) than the SPT sample, and the probability that both samples are drawn from the same distribution is  $< 10^{-5}$  by the Kolmogorov–Smirnov test. A recently published survey<sup>24</sup> of millimetre-identified starbursts

with optical counterparts determined from high-resolution millimetre imaging and redshifts measured from optical spectroscopy or estimated from optical photometry found a median redshift of  $z_{\text{med}} = 2.8$ . Again comparing the distribution of sources at  $z > 2$ , the probability that these objects and the SPT-selected sources are drawn from the same parent distribution is 0.43, indicating rough consistency between our secure redshift determinations and the distribution estimated from the optical methods. A full analysis of the molecular line detections, redshift determinations, residual selection effects, and a



**Figure 3 | The cumulative redshift distribution of luminous dusty starburst galaxies, as measured with different techniques.** The SPT millimetre-selected sample (SPT+ALMA), with redshifts directly determined from spectroscopic observations of the molecular gas in the galaxies, is shown in black. The existing samples of radio-identified (Radio-ID) starbursts<sup>2,14,17,18,27</sup>, with redshifts determined from rest-frame ultraviolet spectroscopy, are compiled in the blue distribution. The redshift distribution<sup>24</sup> of millimetre-identified (mm-ID) starburst galaxies in the COSMOS survey is shown in red/orange, though the majority of redshifts in this sample are derived from optical/infrared photometry of the sources rather than spectroscopy, and therefore less certain. Sources at  $z < 1$  were removed from the previous samples of starburst galaxies to better compare to the selection effect imposed on the SPT sample due to gravitational lensing. The distribution of redshifts for radio-identified sources is incompatible with the distribution for the sample presented in this work. This measurement demonstrates that the fraction of dusty starburst galaxies at high-redshift is greater than previously derived, and that radio-identified samples were biased to lower redshift than the underlying population.

derivation of the intrinsic redshift distribution for the SPT sample is reported elsewhere<sup>23</sup>.

These 26 sources represent less than 25% of the recently completed SPT survey and catalogue. This newly discovered population of high-redshift starbursts will enrich our understanding of obscured star formation in the early Universe. Existing semi-analytic hierarchical models of galaxy evolution<sup>25,26</sup> have already had difficulties reconciling the number of  $z > 4$  systems inferred from previous observational studies<sup>13,24</sup>. The presence of two intensely starbursting galaxies at  $z = 5.7$ , 1 Gyr after the Big Bang, in a sample of just 26 sources, demonstrates that significant reservoirs of dust and molecular gas had been assembled by the end of the epoch of cosmic reionization. As the millimetre-brightest high-redshift starbursts in the sky, the present sample will be key targets for ALMA studies of star-formation physics at high redshift. The gravitational lensing of these sources provides access to diagnostic information from molecular lines that would otherwise take hundreds of times longer to observe, and effective source-plane resolution several times higher than can otherwise be achieved.

Received 28 August 2012; accepted 12 February 2013.

Published online 13 March 2013.

- Lagache, G., Puget, J.-L. & Dole, H. Dusty infrared galaxies: sources of the cosmic infrared background. *Annu. Rev. Astron. Astrophys.* **43**, 727–768 (2005).
- Chapman, S. C., Blain, A. W., Smail, I. & Ivison, R. J. A redshift survey of the submillimeter galaxy population. *Astrophys. J.* **622**, 772–796 (2005).
- Carlstrom, J. E. *et al.* The 10 meter South Pole Telescope. *Publ. Astron. Soc. Pac.* **123**, 568–581 (2011).
- Vieira, J. D. *et al.* Extragalactic millimeter-wave sources in South Pole Telescope survey data: source counts, catalog, and statistics for an 87 square-degree field. *Astrophys. J.* **719**, 763–783 (2010).
- Blain, A. W. & Longair, M. S. Submillimetre cosmology. *Mon. Not. R. Astron. Soc.* **264**, 509–521 (1993).
- Moshir, M., Kopman, G. & Conrow, T. A. O. (eds) *IRAS Faint Source Survey, Explanatory Supplement Version 2* (Infrared Processing and Analysis Center, California Institute of Technology, 1992).
- Bock, D. C.-J., Turtle, A. J. & Green, A. J. A high-resolution radio survey of the Vela supernova remnant. *Astron. J.* **116**, 1886–1896 (1998).
- Blain, A. W. Galaxy-galaxy gravitational lensing in the millimetre/submillimetre waveband. *Mon. Not. R. Astron. Soc.* **283**, 1340–1348 (1996).
- Negrello, M. *et al.* Astrophysical and cosmological information from large-scale submillimetre surveys of extragalactic sources. *Mon. Not. R. Astron. Soc.* **377**, 1557–1568 (2007).
- Hezaveh, Y. D. & Holder, G. P. Effects of strong gravitational lensing on millimeter-wave galaxy number counts. *Astrophys. J.* **734**, 52–59 (2011).
- Hezaveh, Y. D. *et al.* ALMA observations of strongly lensed submillimeter galaxies. *Astrophys. J.* (submitted).
- Ivison, R. J. *et al.* Deep radio imaging of the SCUBA 8-mJy survey fields: submillimetre source identifications and redshift distribution. *Mon. Not. R. Astron. Soc.* **337**, 1–25 (2002).
- Coppin, K. E. K. *et al.* A submillimetre galaxy at  $z = 4.76$  in the LABOCA survey of the Extended Chandra Deep Field-South. *Mon. Not. R. Astron. Soc.* **395**, 1905–1914 (2009).
- Walter, F. *et al.* The intense starburst HDF 850.1 in a galaxy overdensity at  $z \approx 5.2$  in the Hubble Deep Field. *Nature* **486**, 233–236 (2012).
- Solomon, P. M. & Vanden Bout, P. A. Molecular gas at high redshift. *Annu. Rev. Astron. Astrophys.* **43**, 677–725 (2005).
- Capak, P. *et al.* Spectroscopic confirmation of an extreme starburst at redshift 4.547. *Astrophys. J.* **681**, L53–L56 (2008).
- Daddi, E. *et al.* Two bright submillimeter galaxies in a  $z = 4.05$  protocluster in GOODS-North, and accurate radio-infrared photometric redshifts. *Astrophys. J.* **694**, 1517–1538 (2009).
- Daddi, E. *et al.* A CO emission line from the optical and near-IR undetected submillimeter galaxy GN10. *Astrophys. J.* **695**, L176–L180 (2009).
- Riechers, D. A. *et al.* A massive molecular gas reservoir in the  $z = 5.3$  submillimeter galaxy AzTEC-3. *Astrophys. J.* **720**, L131–L136 (2010).
- Cox, P. *et al.* Gas and dust in a submillimeter galaxy at  $z = 4.24$  from the Herschel atlas. *Astrophys. J.* **740**, 63–72 (2011).
- Combes, F. *et al.* A bright  $z = 5.2$  lensed submillimeter galaxy in the field of Abell 773. *HLSJ091828.6+514223. Astron. Astrophys.* **538**, L4–L7 (2012).
- Greve, T. R. *et al.* Submillimeter observations of millimeter bright galaxies discovered by the South Pole Telescope. *Astrophys. J.* **756**, 101–113 (2012).
- Weiβ, A. *et al.* ALMA redshifts of millimeter selected galaxies from the SPT survey. *Astrophys. J.* (in the press).
- Smolcic, V. *et al.* Millimeter imaging of submillimeter galaxies in the COSMOS field: redshift distribution. Preprint at <http://arXiv.org/abs/1205.6470> (2012).
- Baugh, C. M. *et al.* Can the faint submillimetre galaxies be explained in the  $\Lambda$  cold dark matter model? *Mon. Not. R. Astron. Soc.* **356**, 1191–1200 (2005).
- Benson, A. J. GALACTICUS: a semi-analytic model of galaxy formation. *N. Astron.* **17**, 175–197 (2012).
- Banerji, M. *et al.* Luminous starbursts in the redshift desert at  $z \sim 1 - 2$ : star formation rates, masses and evidence for outflows. *Mon. Not. R. Astron. Soc.* **418**, 1071–1088 (2011).

Supplementary Information is available in the online version of the paper.

**Acknowledgements** The SPT is supported by the National Science Foundation, the Kavli Foundation and the Gordon and Betty Moore Foundation. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAJ. The National Radio Astronomy Observatory is a facility of the NSF operated under cooperative agreement by Associated Universities, Inc. Partial support for this work was provided by NASA from the Space Telescope Science Institute. This work is based in part on observations made with Herschel, a European Space Agency Cornerstone Mission with significant participation by NASA. Work at McGill University is supported by NSERC, the CRC programme and ClfAR.

**Author Contributions** J.D.V. and D.P.M. wrote the text. S.C.C. took and reduced optical images and spectroscopy. A.W., C.D.B. and D.P.M. analysed the ALMA spectra. D.P.M., J.S.S. and Y.D.H. analysed the ALMA imaging data. J.D.V. reduced and analysed the Herschel data. Y.D.H. constructed the lens models. C.D.F. reduced optical images. All other authors (listed alphabetically) have contributed as part of the South Pole Telescope collaboration, by their involvement with the construction of the instrument, the initial discovery of the sources, multi-wavelength follow-up, and/or contributions to the text.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.D.V. ([vieira@caltech.edu](mailto:vieira@caltech.edu)).



# A multi-directional backlight for a wide-angle, glasses-free three-dimensional display

David Fattal<sup>1</sup>, Zhen Peng<sup>1</sup>, Tho Tran<sup>1</sup>, Sonny Vo<sup>1</sup>, Marco Fiorentino<sup>1</sup>, Jim Brug<sup>1</sup> & Raymond G. Beausoleil<sup>1</sup>

Multiview three-dimensional (3D) displays can project the correct perspectives of a 3D image in many spatial directions simultaneously<sup>1–4</sup>. They provide a 3D stereoscopic experience to many viewers at the same time with full motion parallax and do not require special glasses or eye tracking. None of the leading multiview 3D solutions is particularly well suited to mobile devices (watches, mobile phones or tablets), which require the combination of a thin, portable form factor, a high spatial resolution and a wide full-parallax view zone (for short viewing distance from potentially steep angles). Here we introduce a multi-directional diffractive backlight technology that permits the rendering of high-resolution, full-parallax 3D images in a very wide view zone (up to 180 degrees in principle) at an observation distance of up to a metre. The key to our design is a guided-wave illumination technique based on light-emitting diodes that produces wide-angle multiview images in colour from a thin planar transparent light-guide. Pixels associated with different views or colours are spatially multiplexed and can be independently addressed and modulated at video rate using an external shutter plane. To illustrate the capabilities of this technology, we use simple ink masks or a high-resolution commercial liquid-crystal display unit to demonstrate passive and active (30 frames per second) modulation of a 64-view backlight, producing 3D images with a spatial resolution of 88 pixels per inch and full-motion parallax in an unprecedented view zone of 90 degrees. We also present several transparent hand-held prototypes showing animated sequences of up to six different 200-view images at a resolution of 127 pixels per inch.

Ideally, a perfect 3D display would reproduce the set of all light rays (or lightfield<sup>5</sup>) from a 3D scene. Although standard holography can perform this task very well, the recording of a holographic medium is too slow to permit real-time operation<sup>6</sup>. Autostereoscopic multiview 3D displays are ‘the next-best thing’: they discretize the lightfield into narrowly spaced views to create the illusion of continuous parallax and a 3D stereoscopic effect up to a certain viewing distance<sup>1,2</sup>. They can be realized using pure geometrical optics techniques such as multi-projector<sup>7</sup>, parallax barrier<sup>3</sup>, integral imaging<sup>8</sup> or a combination of these<sup>4,9</sup>. Multi-projector solutions have demonstrated impressively large 3D images with many views and high resolution, but they are difficult to implement on a mobile device. Recent lenticular-based solutions (see <http://www.alioscopy.com> and <http://www.dimencodisplays.com>) have seemed promising for TV applications owing to their compatibility with standard liquid-crystal displays (LCDs); however, they provide only a limited amount of head movement (with the sets of views repeating several times across the view zone), and low resolution.

Diffractive optics offers another way of creating a multiview lightfield. Digital holographic displays use a spatial light modulator to update small holographic cells (sometimes called hogels<sup>10</sup>), each generating a few light rays. Despite recent progress (<http://www.zebraimaging.com>), the required density of active pixels limits the field of view (FOV) and precludes the operation of the digital holographic displays at video rate. Using a static fringe pattern with an external modulator<sup>11</sup> can solve the video rate problem for a more modest number of views. Still, because

those diffractive approaches rely on the first-order diffraction of a highly collimated free-space light source, they are hard to integrate, have limited power efficiency, and are plagued by the presence of specular light and unwanted diffraction orders (even for complex diffractive elements such as blazed or sub-wavelength gratings<sup>12</sup>).

Here we introduce a diffractive backlight solution that seems well suited to realizing a multiview 3D display for mobile devices. It features (1) full motion parallax in a wide view zone (90° FOV demonstrated), (2) static ‘directional pixels’ (binary diffraction gratings) that can be manufactured at low cost and individually modulated at video rate, and (3) edge lighting by standard light-emitting diode (LED) light with a compact collimation system. In addition, it can render colour without the need for colour filters, allowing full transparency of the display (with transparent electronics).

The cornerstone of our display architecture is a set of directional grating pixels shown schematically in Fig. 1. The gratings are etched or deposited on the backlight surface, and illuminated by collimated light that is guided in the backlight substrate by total internal reflection. As a result of first-order diffraction, light with input planar momentum  $k_{\text{in}} = (n_{\text{eff}}, 0, 0)2\pi/\lambda$  is scattered out of the backlight in a well-defined direction characterized by the normalized output vector  $k_{\text{out}} = (n_x, n_y, n_z)2\pi/\lambda$ , where:

$$n_x = n_{\text{eff}} - (\lambda/A)\cos\phi \quad (1)$$

$$n_y = -(\lambda/A)\sin\phi \quad (2)$$

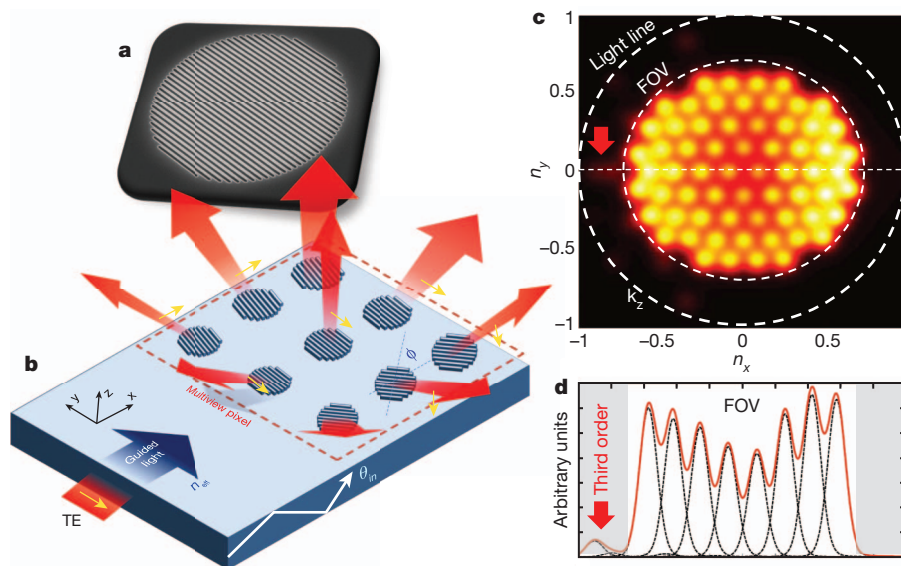
Also,  $n_x^2 + n_y^2 + n_z^2 = 1$ ,  $n_{\text{eff}}$  is the effective index of propagation of the input light along the  $x$  axis,  $A$  is the grating pitch and  $\phi$  is the groove orientation with respect to the  $y$  axis. A group of directional gratings covering all views in the FOV forms a multiview pixel—the basic unit cell of our backlight.

The guided nature of the input light ( $n_{\text{eff}} > 1$ ) completely suppresses specular light in the FOV. It also repels higher diffraction orders away from the display normal, ensuring a wide view zone free of ghost images. In the Supplementary Information, we prove that the extent of the ‘clean’ view zone (measured in terms of the angle  $\alpha$  from the vertical axis ( $z$  axis) is:

$$\sin\alpha = n_{\text{eff}}/2 \quad (3)$$

For a glass or plastic backlight, this corresponds to a FOV ( $2\alpha$ ) of 90° (and a high-refractive-index glass<sup>13</sup> could be used to reach an FOV close to 180°). The guided light illumination is essential for another reason: it allows spatial multiplexing of three pixel sets that can be selectively addressed by changing the illumination angle. This important feature allows the spatial multiplexing of colour without using colour filters, potentially increasing the power efficiency and enabling a completely transparent architecture. This illumination technique

<sup>1</sup>Hewlett-Packard Laboratories, 1501 Page Mill Road, Palo Alto, California 94304, USA.



**Figure 1 | Multi-directional backlight concept.** **a**, Scanning electron micrograph of grating pixel fabricated by conventional photolithography. **b**, Schematic of the multi-directional backlight under collimated light illumination. ‘Directional’ grating pixels scatter the input light into the view zone of the display by first-order diffraction. Owing to the large planar momentum of the input light, unwanted diffraction orders fall far from the display normal. Under illumination by TE (transverse electric) polarized light, the output polarization is well preserved, which facilitates the external modulation of the backlight using commercial liquid-crystal technology.  $\theta_{in}$ ,

the angle of propagation of the input light in the backlight, with respect to the horizontal plane. **c**, Full-wave simulation of the radiation pattern from a 64-view backlight with 12- $\mu\text{m}$ -diameter red grating pixels under illumination by a collimated LED light source (20 nm spectral width,  $\pm 3.5^\circ$  collimation approximating our experimental lighting conditions). The upward flux is  $k_z|E|^2$ . **d**, Cross-section of **c** along the horizontal axis with individual view contributions (dotted black lines). The diffraction efficiency for each view is about 5% for 100-nm-deep grating grooves. The red arrow indicates the presence of a third-order diffracted beam outside the FOV.

also allows us to create animated sequences of 3D images in a very simple way.

The use of diffractive pixels allows us to define almost arbitrary spatio-angular characteristics for the emitted lightfield. For example, we can increase the angular density near the normal direction and decrease it at sharper view angles. We can increase the number of horizontal views and reduce the number of vertical views. We can adopt a hexagonal angular grid (as in Fig. 1c) to obtain more uniform transitions between views. Unlike most solutions based on geometrical optics<sup>8</sup>, we have complete freedom to define the overlap between neighbouring views, allowing us to control the amount of between-view leakage and angular aliasing<sup>14,15</sup>. More importantly, it permits us to mix standard, non-directional pixels (implemented as chirped broadband gratings) with directional pixels on the backlight surface to enable high-definition two-dimensional 2D content and lower-resolution 3D content simultaneously.

The visual experience provided by a multiview 3D display can be characterized by its spatial resolution (multiview pixel size  $p$ ), and angular resolution (between-view distance  $\Delta\theta$ ), or number of views  $N = (2\sin\alpha/\Delta\theta)^2$ . The maximum distance at which the stereo effect can be continuously perceived in the FOV is  $z_{3D} \approx 6.3\text{ cm}/\Delta\theta$ , where 6.3 cm is the average human eye separation<sup>16</sup>. The depth of field ( $p/\Delta\theta$ ) measures the maximum altitude of an image that can be displayed at the native resolution  $p$  of the display with a disparity of less than one pixel between neighbouring views (larger disparity content will show

some blurring and ‘jumps’ between views<sup>15</sup>). The effective resolution at height  $z$  (where  $z = 0$  is the surface of the display) is  $p_{\text{eff}}(z) = \max(p, z\Delta\theta)$ . The maximum number of views that our backlight can accommodate is:

$$N \approx p\sin\alpha/(\lambda\beta/n) \quad (4)$$

where  $n$  is the number of spatially multiplexed pixel sets ( $n = 3$  for a red–green–blue (RGB) backlight) and  $\beta$  is a geometrical factor that depends on the spatio-angular pixel layout ( $\beta = 1.22$  for a close-packed square array of circular pixels, with a hexagonal distribution of views within the FOV and between-view separation corresponding to one Rayleigh width). Table 1 illustrates the imaging performance of a full-colour backlight with  $90^\circ$  FOV for various pixel sizes. The backlight can be improved by incorporating a set of ‘broadband’ pixels to increase the resolution of low-disparity content without decreasing the depth of field, or by using time-multiplexing schemes to increase the effective view number. Such refinements will be the topic of future study.

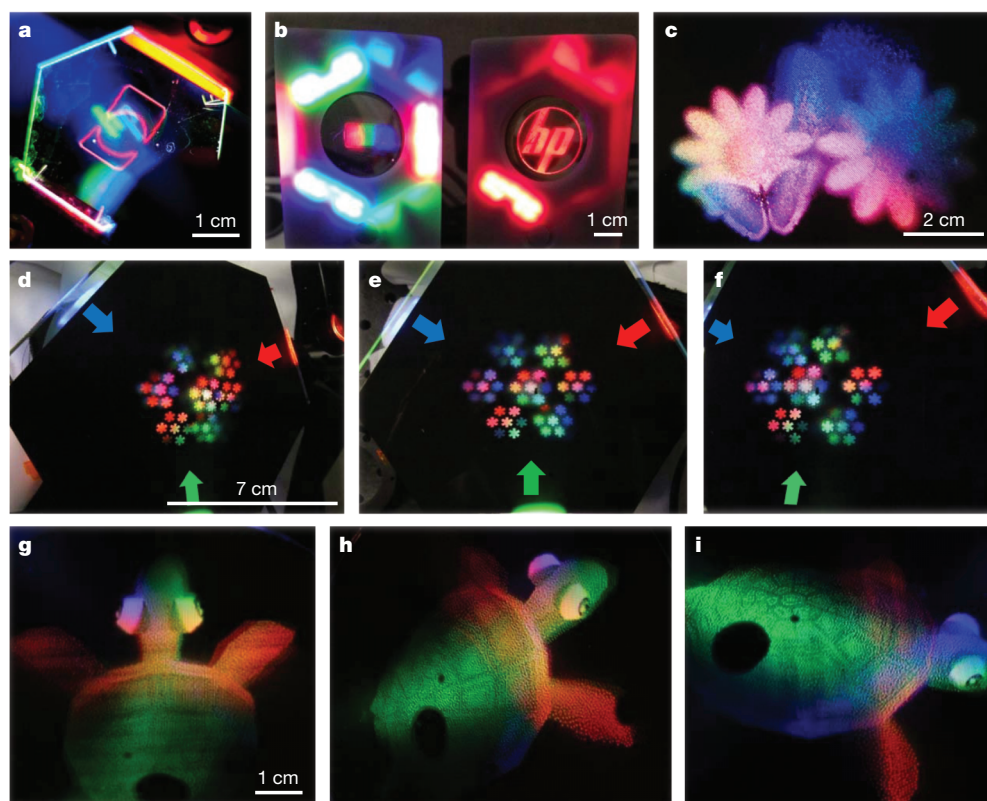
In practice, the angular divergence of the views is broadened beyond the diffraction limit by the finite spectral and angular distribution of the input light. Figure 1c and d presents the simulated radiation pattern<sup>17</sup> of our 64-view backlight prototype, designed with an FOV of  $90^\circ$  and inter-view separation  $\theta$  of  $10^\circ$  to accommodate an input-beam spectral width of 20 nm and angular width of  $\pm 3.5^\circ$ . The brightness variation across the FOV (centred around a diffraction efficiency of 5%

**Table 1 | Theoretical characteristics of diffraction-limited backlight**

3D imaging characteristics of a full-colour backlight	Diffraction-limited performance*			64-view experimental backlight of Fig. 1c and d
Multiview pixel size, $p$ (mm)	0.25	0.5	1.0	0.288
Number of views per colour, $N$	133	266	531	64
Angular separation between views, $\Delta\theta$ ( $^\circ$ )	7.0	5.0	3.5	10
Diameter of each grating sub-pixel, $D$ ( $\mu\text{m}$ )	12.5	17.7	25.1	12
Stereo distance, $z_{3D}$ (cm)	51	73	103	36
Depth of field (mm)	2.0	5.8	16.3	1.7
Effective resolution at height 2 cm, $p_{\text{eff}}$ (mm)	2.5	1.7	1.2	3.5

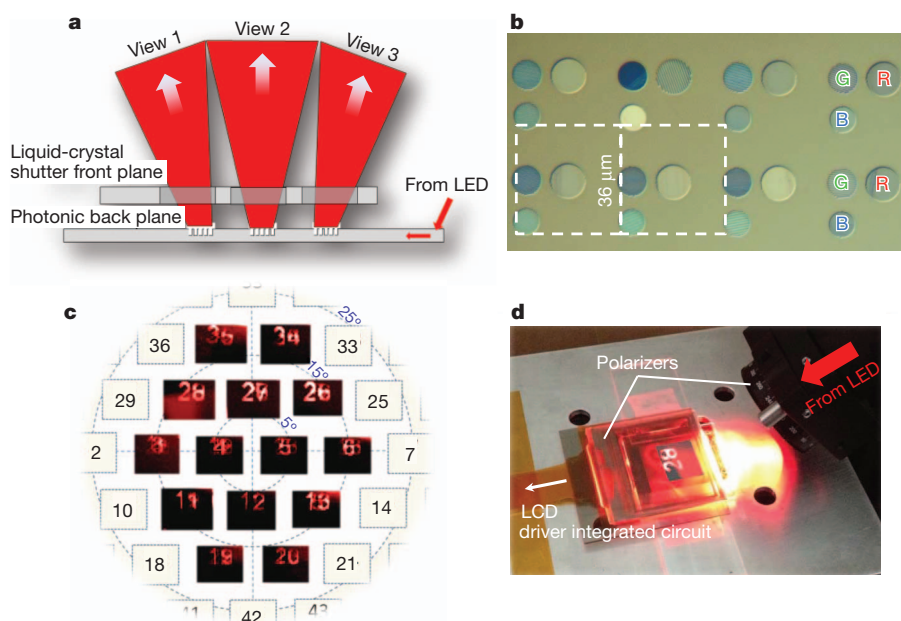
\* Shown are three distinct designs with increasing pixel size, highlighting the trade-off between spatial resolution (pixel size  $p$ ) and angular resolution (angular separation  $\Delta\theta$ ).





**Figure 2 | 3D images from a passively modulated backlight.** **a, b**, Portable, transparent backlight prototype, with resolution 127 pixels per inch, 200 views, wide view zone of  $90^\circ$  (full parallax). The red HP logo in **b** is animated in a 'breathing' motion by periodically cycling through one of six LED arrays (see Supplementary Video). **c**, Colour dithered 3D image obtained by masking dark pixels from a fully covered backlight, using a high-resolution ink-printed

pattern. **d–f**, Steep angle views from the same backlight as **c** with a test pattern (colour 'snow flakes') showing 4 cm of rendered depth and colour mixing. Scale bar in **d** refers to **e** and **f** as well; scale bar in **g** refers to **h** and **i** as well. We note that the elements of the image with highest disparity (located at  $z = 2$  cm) show some blurring, owing to the engineered view overlap of the backlight. **g–i**, Turtle 3D image viewed from various viewpoints.



**Figure 3 | Actively modulated backlight.** **a**, Schematics of a side-illuminated backlight actively modulated by an LCD shutter plane. The shutter plane must be located close enough to the backlight surface to avoid beam walk-off. **b**, Optical microscope image of the pixel arrangement on the backlight surface. Although all three RGB pixels were present on the backlight, we used only the red pixels for that prototype. The dotted white squares indicate the alignment of the liquid-crystal cell with respect to the pixel array. **c**, Multiview image

showing a different number in different view zones. Images corresponding to neighbouring views 'bleed' into each other owing to the designed view overlap, introduced to smooth out between-view transitions. **d**, Global view of the LCD modulated prototype. Red light from an external LED is expanded and collimated using free-space optics elements, and linearly polarized before it enters the backlight directly from a polished edge. The LCD shutter plane was thinned to  $50\ \mu\text{m}$  and aligned to the pixel array with an accuracy of about  $2\ \mu\text{m}$ .

for light diffracted above the grating, for 100-nm grating grooves) and the predicted view divergence (about  $6.5^\circ$  full width at half maximum, FWHM) compare well with the data reported in the Supplementary Information.

We present a series of prototypes of increased complexity to showcase the features of our backlight technology. We fabricate hand-held devices featuring a transparent backlight with partial pixel coverage, able to project static 3D RGB images or animated sequences of up to six monochrome 3D images ( $90^\circ$  FOV, 200  $\mu\text{m}$  resolution, 200 views per colour; see Fig. 2a and b). We also fabricate several 6-inch backlights with full pixel coverage (see Fig. 1c and d), on which we overlaid high-resolution binary masking patterns (ink on plastic from a high-resolution printer) to create larger 64-view images with colour mixing and dithering (Fig. 2c–i).

Finally, we use the front plane of a Sony LCD pico-projector to modulate a 1-inch 64-view backlight at a rate of 30 frames per second. As shown in Fig. 3, the backlight contains a 720 by 720 array of grating pixels, with a pitch of 36  $\mu\text{m}$ . Because the gratings in the backlight do not significantly change the polarization of horizontally (transverse electric) polarized light, only the output polarizer is needed for the LCD. Figure 3d shows a multiview image created by the LCD in which each view displays a different number from 1 to 64. To avoid ‘beam walk-off’ from the grating pixels to their respective liquid-crystal cell, the top glass cover needs to be thinned to about 20  $\mu\text{m}$ . The resulting devices were not strong enough for extensive testing, so here we report the operation of a stable device featuring a LCD cover thickness of about 50  $\mu\text{m}$ , allowing the modulation of 14 of the original 64 views (the unmodulated views are greyed out on Fig. 3d). The pictures taken of the 14 central views show a small amount of between-view mixing, in qualitative agreement with the simulation of Fig. 1c and d. At the centre of views 12 or 28, for instance, no leakage from neighbouring views is discernible. In view 27, the view point leans towards view 26 and the between-view leakage is apparent. See the Supplementary Information for videos of all static and active demonstrations.

Our multi-directional backlight provides the basis for a very efficient display, because it does not require colour filters, and can use multiple passes of the incident light, much as in a standard LCD. In its present form, the intensity of the input light decreases from the edge to the centre, which can be compensated in a variety of ways (for example, by adjusting scattering strength via groove depth or area variation). We are currently working on other pixel modulation methods that do not require polarized light and should increase the efficiency even further.

We believe that the unique combination of wide-angle 3D performance, high spatial resolution, ease of modulation at video rate, compact form factor and low manufacturing cost make our multi-directional backlight technology very promising for multiview 3D mobile display applications.

## METHODS SUMMARY

The glass backlight prototypes presented in this paper are made of Schott glass B270, covered by a 100-nm layer of silicon nitride fabricated by plasma-enhanced chemical vapour deposition. Some of the static images were defined by electron-beam lithography using a 200-nm PMMA photoresist layer covered by 2 nm of

evaporated chrome. The nitride layer was etched in an Oxford reactive ion etching system using chlorine chemistry. The  $720 \times 720$  grating pixel array for the LCD modulated backlight was defined by a 193-nm deep-ultraviolet photolithography process. The finite-difference time domain simulations were performed on a Hewlett-Packard-owned computer cluster using the parallel version of MEEP<sup>17</sup> with 256 simultaneous cores.

Received 12 October 2012; accepted 30 January 2013.

1. Dodgson, N. Autostereoscopic 3D displays. *Computer* **38**, 31–36 (2005).
2. Okoshi, T. *Three-Dimensional Imaging Techniques* (Atara Press, 2011).
3. Wetzstein, G., Lanman, D., Heidrich, W. & Raskar, R. Layered 3D: tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.* **30** (95), 1–12 (2011).
4. Wetzstein, G., Lanman, D., Hirsch, M. & Raskar, R. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph. (Proc. SIGGRAPH)* **31**, 1–11 (2012).
5. Levoy, M. & Hanrahan, P. Light field rendering. In *Proc. 23rd Ann. Conf. on Computer Graphics And Interactive Techniques (SIGGRAPH '96)* 31–42 (ACM, New York, 1996).
6. Blanche, P. A. et al. Holographic three-dimensional telepresence using large-area photorefractive polymer. *Nature* **468**, 80–83 (2010).
7. Kawakita, M. et al. 3D image quality of 200-inch glasses-free 3D display system. *Proc. SPIE* **8288**, <http://dx.doi.org/10.1117/12.912274> (2012).
8. van Berkel, C. & Clarke, J. A. Characterization and optimization of 3D-LCD module design. *Proc. SPIE* **3012**, 179–186 (1997).
9. Takaki, Y. & Nago, N. Multi-projection of lenticular displays to construct a 256-view super multi-view display. *Opt. Express* **18**, 8824–8835 (2010).
10. Lucente, M. *Diffraction-Specific Fringe Computation for Electro-Holography*. PhD thesis (MIT, 1994).
11. Kulick, J. H. et al. Partial pixels: a three-dimensional diffractive display architecture. *J. Opt. Soc. Am. A* **12**, 73–83 (1995).
12. Astilean, S., Lalanne, P., Chavel, P., Cambri, E. & Launois, H. High-efficiency sub-wavelength diffractive element patterned in a high-refractive-index material for 633 nm. *Opt. Lett.* **23**, 552–554 (1998).
13. Fu, J. Optical glass. US patent 8,187 986 (2012).
14. Ramachandra, V., Hirakawa, K., Zwicker, M. & Nguyen, T. Spatioangular prefiltering for multiview 3D displays. *IEEE Trans. Vis. Comput. Graph.* **17**, 642–654 (2011).
15. Zwicker, M., Matusik, W., Durand, F. & Pfister, H. Antialiasing for automultiscopic 3D displays. In *Eurographics Symposium on Rendering* (2006).
16. Dodgson, N. A. Variation and extrema of human interpupillary distance. *Proc. SPIE* **5291**, 36–46 (2004).
17. Oskooi, A. F. et al. MEEP: a flexible free-software package for electromagnetic simulations by the FDTD method. *Comp. Phys. Commun.* **181**, 687–702 (2010).

Supplementary Information is available in the online version of the paper.

**Acknowledgements** We are grateful to C. Santori for many suggestions on the backlight design, A. Jeans for performing the roll-to-roll imprint jobs, A. Said for generating some of the multiview images shown in Fig. 2, L. Kiyama and W. Mack for the printed circuit board design of the portable prototypes, R. Cobene for his help with packaging, P. Beck for her help in the cleanroom, and B. Culbertson and C. Patel for their support of the project.

**Author Contributions** All authors contributed extensively to the work presented in this paper. D.F. conceived the multi-directional backlight concept. D.F. and J.B. led the technical effort to realize the various prototypes. D.F., Z.P., S.V. and T.T. were responsible for the nanometre-scale fabrication of the backlight. J.B. and M.F. were responsible for the optical design and assembly of the prototypes. M.F. designed the illumination systems and part of the other electronics for the prototypes. R.G.B. supervised and coordinated the project. All authors contributed to the data analysis. D.F. and R.G.B. prepared the manuscript with input from J.B., M.F. and Z.P.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.F. ([david.fattal@hp.com](mailto:david.fattal@hp.com)).



# Changes in global nitrogen cycling during the Holocene epoch

Kendra K. McLauchlan<sup>1</sup>, Joseph J. Williams<sup>1</sup>, Joseph M. Craine<sup>2</sup> & Elizabeth S. Jeffers<sup>3</sup>

Human activities have doubled the pre-industrial supply of reactive nitrogen on Earth, and future rates of increase are expected to accelerate<sup>1</sup>. Yet little is known about the capacity of the biosphere to buffer increased nitrogen influx. Past changes in global ecosystems following deglaciation at the end of the Pleistocene epoch provide an opportunity to understand better how nitrogen cycling in the terrestrial biosphere responded to changes in carbon cycling. We analysed published records of stable nitrogen isotopic values ( $\delta^{15}\text{N}$ ) in sediments from 86 lakes on six continents. Here we show that the value of sedimentary  $\delta^{15}\text{N}$  declined from 15,000 years before present to  $7,056 \pm 597$  years before present, a period of increasing atmospheric carbon dioxide concentrations and terrestrial carbon accumulation<sup>2</sup>. Comparison of the nitrogen isotope record with concomitant carbon accumulation on land and nitrous oxide in the atmosphere suggests millennia of declining nitrogen availability in terrestrial ecosystems during the Pleistocene–Holocene transition around 11,000 years before present. In contrast, we do not observe a consistent change in global sedimentary  $\delta^{15}\text{N}$  values during the past 500 years, despite the potential effects of changing temperature and nitrogen influx from anthropogenic sources. We propose that the lack of a single response may indicate that modern increases in atmospheric carbon dioxide and net carbon sequestration in the biosphere have the potential to offset recent increased supplies of reactive nitrogen in some ecosystems.

Humans have drastically altered the global nitrogen (N) cycle by doubling rates of biotic and abiotic  $\text{N}_2$  fixation. Consequently, supplies of reactive N to many terrestrial ecosystems have increased<sup>3</sup>, and are projected to continue to increase globally by 10–15% within 25 yr (ref. 4). This increase in reactive N has led to suggestions that Earth has crossed a critical boundary regarding the N cycle<sup>5</sup>. However, despite the certainty of increased supply rates, it is unclear whether N availability—the supply of N relative to demand by plants and microbes—has increased on a global scale. Uncertainty about changes in net ecosystem response to increased reactive N hinders predictions of the future function of important ecosystem processes, such as primary productivity, decomposition and nitrate export to waters. Recent measurements of N availability on sub-annual to decadal timescales have demonstrated conflicting patterns. Although some areas of high deposition have shown signs of increased N availability<sup>6</sup>, many terrestrial ecosystems have shown a surprising capacity to retain added reactive N without evidence of adverse effects, and N availability has apparently declined in other ecosystems during the past few decades<sup>7</sup>.

Given the uncertainty about changes in terrestrial N availability, extending records of N cycling before human manipulation of the N cycle would be useful. A variety of empirical correlations<sup>8</sup> and theoretical models<sup>9</sup> indicate that the variation in past N cycling can be captured in the  $\delta^{15}\text{N}$  values of leaves<sup>10</sup>, wood<sup>11</sup>, soils<sup>12</sup> and sediments<sup>13</sup>. Of these, lacustrine sediments provide the longest-term reconstructions and integrate over the largest area, although they also integrate over the greatest number of processes that can affect the ultimate isotopic values. The basic premise is that  $\delta^{15}\text{N}$  measured in organic

material integrates fractionation processes, such that  $\delta^{15}\text{N}$  values provide an index of the processes in the N cycle that discriminate against  $^{15}\text{N}$ , particularly gaseous loss pathways<sup>14</sup>. When N supply is high relative to biotic demand, N is lost through fractionating pathways and the remaining ecosystem N is enriched in  $^{15}\text{N}$ . In the few studies that have measured several ecosystem compartments, values of terrestrial  $\delta^{15}\text{N}$ , stream  $\delta^{15}\text{N}$  in  $\text{NO}_3^-$ , and lacustrine sedimentary  $\delta^{15}\text{N}$  were correlated<sup>14,15</sup>.

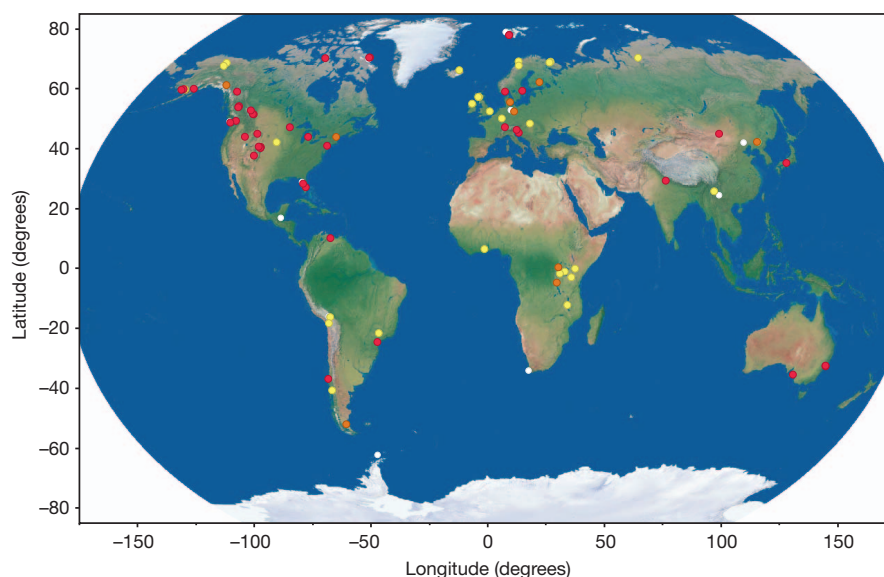
To better understand the patterns of N cycling over century to multi-millennial timescales, we compiled published records of sedimentary  $\delta^{15}\text{N}$  from 86 lakes distributed on six continents to assess changes in ecosystem-level N cycling over the past 15,000 yr (Fig. 1). This time period encompasses the Pleistocene–Holocene transition—a period of rapid warming, atmospheric  $\text{CO}_2$  increases, and primary succession in newly deglaciated landscapes—and also includes the past 500 yr during which there have been unprecedented anthropogenic global changes. For inclusion in this analysis, sites had to meet minimum criteria for sampling density and duration of sedimentary record, leading to sample sizes of 37 sites for the analysis from 15,000 years before present, and 58 sites for the 500-year analysis. The original chronology of each study was reassessed and expressed in calendar years before present (cal. yr BP), where present (0 cal. yr BP) is defined as AD 1950. Sites are distributed broadly over the globe, representing a variety of climates, biomes and degrees of human impact (Fig. 1, Supplementary Fig. 3). Modern mean annual precipitation at these sites ranges from 118 to 2,516 mm and mean annual temperature ranges from  $-17$  to  $26^\circ\text{C}$ .

During the Pleistocene–Holocene transition when global terrestrial ecosystems were accumulating C, sedimentary  $\delta^{15}\text{N}$  declined globally for approximately 8,000 yr (Fig. 2a). From 15,000 cal. yr BP to an inflection point at  $7,056 \pm 597$  cal. yr BP, sedimentary  $\delta^{15}\text{N}$  declined at a rate of  $0.25 \pm 0.03\text{‰ kyr}^{-1}$  ( $P < 0.001$ ). Of the 17 sites with declining trends during this period, 11 were significant, whereas only 3 of the 10 increasing trends were significant. The sites with declining records were located in both tropical and temperate areas and both hemispheres; some had been glaciated and some had not. Since the minimum value at 7,050 cal. yr BP, global sedimentary  $\delta^{15}\text{N}$  has been increasing at a gradual rate of  $0.08 \pm 0.04\text{‰ kyr}^{-1}$  ( $P < 0.05$ ). This is 32% of the rate of decline during the previous 8,000 yr.

Despite the general trend for increasing sedimentary  $\delta^{15}\text{N}$  over the past 7,000 yr, high-resolution modern sedimentary  $\delta^{15}\text{N}$  records show no evidence of consistent changes in global N cycling during the past 500 yr (Fig. 2b). Analysis of a larger set of lakes with higher sampling density, during the time window from AD 1500 to AD 2005, shows that sedimentary  $\delta^{15}\text{N}$  has not been significantly changing during recent times ( $n = 58$ ,  $P = 0.2$ ). Only 28 of the 58 sites in the analysis of the past 500 yr have sedimentary  $\delta^{15}\text{N}$  records that show a significant increase or decrease over time. Sedimentary  $\delta^{15}\text{N}$  has significantly increased at 10 sites and significantly decreased at 18 sites since AD 1500.

Regression analysis suggests that both climatic and anthropogenic variables affected the direction of change in sedimentary  $\delta^{15}\text{N}$

<sup>1</sup>Department of Geography, Kansas State University, Manhattan, Kansas 66506, USA. <sup>2</sup>Division of Biology, Kansas State University, Manhattan, Kansas 66506, USA. <sup>3</sup>Biodiversity Institute, Department of Zoology, University of Oxford, Oxford OX1 3PS, UK.



**Figure 1 | Locations of the 86 sites analysed for lacustrine sedimentary  $\delta^{15}\text{N}$  in this study.** Red indicates a sedimentary record that was analysed only for the past 500 yr, yellow indicates a record used for the 15,000 cal. yr BP analysis, and

orange indicates a site that met analytical criteria for both time periods. Sites with published sedimentary  $\delta^{15}\text{N}$  records that were not included in this study are coloured white. Full site information is available in Supplementary Table 1.

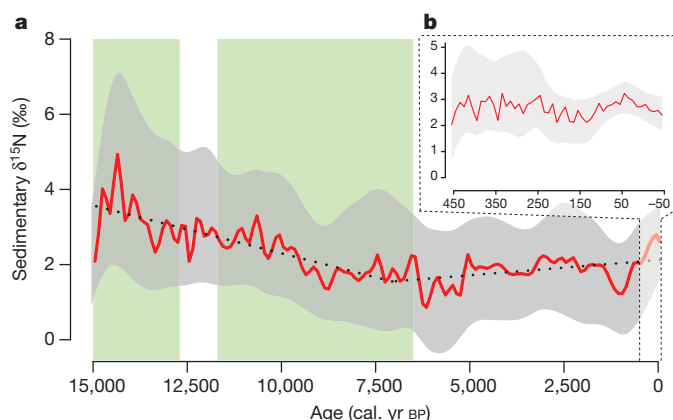
at individual sites during the past 500 yr (Fig. 3). Overall, lakes from warm, low-elevation regions are more likely to have increasing values of sedimentary  $\delta^{15}\text{N}$  towards the present, while those from cold, high-elevation regions are more likely to have decreasing values of sedimentary  $\delta^{15}\text{N}$  towards the present. For example, the sediments from a typical lake at 3,000 m where mean annual temperature was  $-10^\circ\text{C}$  would have decreased in  $\delta^{15}\text{N}$  by 1.3‰ per century for the past 500 yr, while sediment from lakes at sea level where mean annual temperature was  $25^\circ\text{C}$  would have increased by 1.5‰ per century over the same time period. Anthropogenic factors also influence observed trends of sedimentary  $\delta^{15}\text{N}$ . Lakes in regions with a large human footprint—as calculated from population density, land transformation, accessibility and electrical power infrastructure<sup>16</sup>—are more likely to have sedimentary  $\delta^{15}\text{N}$  that decreased towards the present.

The global patterns of sedimentary  $\delta^{15}\text{N}$  during the past 15,000 yr are consistent with a strong influence of terrestrial C sequestration on the global N cycle. The prolonged retreat of continental ice sheets in

the Northern Hemisphere during the Pleistocene–Holocene transition exposed approximately  $5 \times 10^{12} \text{ m}^2$  of land area<sup>17</sup> that accumulated C and N after colonization by plants. During the early Holocene,  $\sim 600 \text{ Pg}$  of C were sequestered in terrestrial ecosystems<sup>17</sup>, with strong net terrestrial C accumulation from 15,000 to 7,000 cal. yr BP (ref. 2). The N budgets that accompanied this global C sequestration are almost completely unknown<sup>17</sup>, although several local- or regional-scale studies provide estimates of rapid rates of N accumulation associated with primary succession<sup>18,19</sup>. On the global scale, if the additional organic matter had a C:N ratio of 100 (ref. 14) it would have required 6,000 Tg of N to accumulate the organic matter from the Last Glacial Maximum to pre-industrial Holocene.

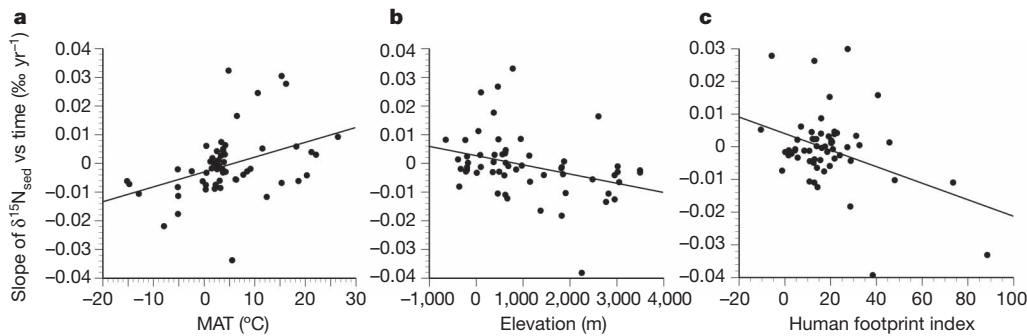
Two underlying mechanisms associated with terrestrial C sequestration could be contributing to the broad declines in sedimentary  $\delta^{15}\text{N}$  seen from 15,000 to 7,000 cal. yr BP (Fig. 4). First, increased  $\text{N}_2$  fixation would add N to local terrestrial ecosystems with an isotopic signature initially near 0‰, thereby potentially reducing sedimentary  $\delta^{15}\text{N}$  values<sup>20</sup> despite increasing terrestrial N supplies. Under equilibrium conditions, the addition of fixed N often leads to increases in ecosystem  $\delta^{15}\text{N}$  by increasing N supply and fractionating losses<sup>21</sup>; however, if ecosystems are accumulating and sequestering C then  $\text{N}_2$  fixation can reduce ecosystem  $\delta^{15}\text{N}$  on century to millennial timescales<sup>22</sup>. Whether  $\text{N}_2$ -fixation during the initial stages of primary succession after deglaciation increased or decreased ecosystem  $\delta^{15}\text{N}$  at individual sites is an open question; however, we found that sites which were not glaciated during the Last Glacial Maximum also exhibited declines in sedimentary  $\delta^{15}\text{N}$  from 15,000 to 7,000 cal. yr BP, thus additional mechanisms were probably driving the decline in  $\delta^{15}\text{N}$  values.

Alternatively, C sequestration during the early Holocene could have reduced local terrestrial N availability, which would be reflected in lower values of ecosystem  $\delta^{15}\text{N}$  (that is, foliage and soil) and of the N transported to lakes. The proposed hypothesis that the  $\delta^{15}\text{N}$  of lacustrine sediments reflects reduced N availability is further supported by observed declines in atmospheric concentrations of  $\text{N}_2\text{O}$  during the early Holocene (Supplementary Fig. 7).  $\text{N}_2\text{O}$  is produced by the fractionating gaseous loss pathways of nitrification and denitrification in soils with high N supply relative to demand. The Holocene minimum of atmospheric  $\text{N}_2\text{O}$  occurs between 8,000 and 6,000 cal. yr BP (ref. 23; Supplementary Fig. 7), temporally coincident with the minimum in sedimentary  $\delta^{15}\text{N}$ , and it has potentially been linked to reduced terrestrial  $\text{N}_2\text{O}$  production in soils.



**Figure 2 | Changes in lacustrine sedimentary  $\delta^{15}\text{N}$  during the late Pleistocene and Holocene.** **a**, A smoothing spline curve (0.05 smoothing parameter) fitted to the means of sites in 100-yr bins is shown (red) with 95% bootstrapped confidence intervals (grey). Declines in sedimentary  $\delta^{15}\text{N}$  from 15,000 cal. yr BP to the breakpoint at  $7,056 \pm 597$  cal. yr BP correspond with periods of global net terrestrial carbon gain (shaded green)<sup>4</sup>. Dotted black line is the breakpoint regression. **b**, A different set of high-resolution sedimentary  $\delta^{15}\text{N}$  records shows no net change over the past 500 yr.





**Figure 3 | Three local environmental variables explain site-specific trajectories of sedimentary  $\delta^{15}\text{N}$  in the past 500 yr.** a–c, Partial residual plots are shown between mean annual temperature (MAT; partial  $r^2 = 0.11$ ,  $P < 0.005$ ; a), elevation (partial  $r^2 = 0.08$ ,  $P < 0.01$ ; b) and a human footprint index (partial  $r^2 = 0.1$ ,  $P < 0.005$ ; c) and the slope of site-specific regressions

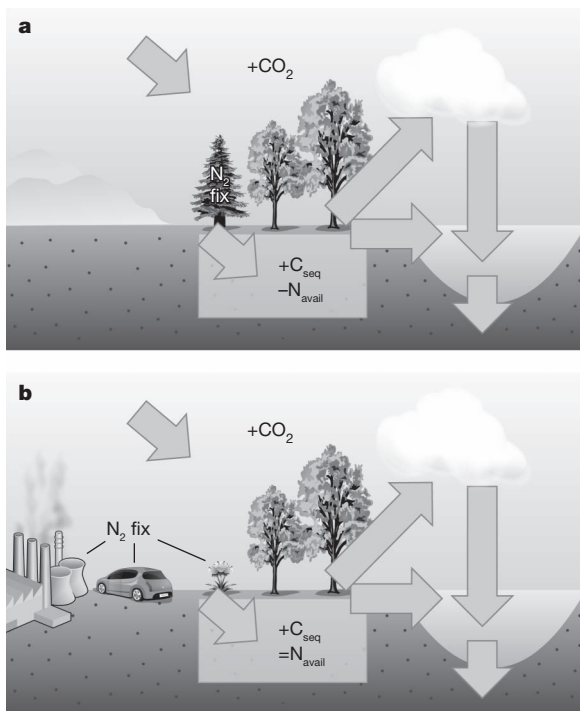
Any reduction in N availability due to net C sequestration probably would have been amplified by increasing atmospheric  $\text{CO}_2$  concentrations during the Pleistocene–Holocene transition. Atmospheric  $\text{CO}_2$  concentrations increased from 188 p.p.m.v. at ~15,000 cal. yr BP to 260 p.p.m.v. by 10,000 cal. yr BP (ref. 24). This sustained increase in atmospheric  $\text{CO}_2$  could have generated progressive nitrogen limitation<sup>25</sup> in plants as their demand for N increased with the higher levels of productivity and N was sequestered (along with C) in plant biomass and soil organic matter pools. This could have contributed to lowered N availability in terrestrial ecosystems for thousands of years. The effects of local increases in  $\text{N}_2$  fixation and reduced N availability are not mutually exclusive, and in fact the processes are linked. Low N

between sedimentary  $\delta^{15}\text{N}$  ( $\delta^{15}\text{N}_{\text{sed}}$ ) and cal. yr BP for the past 500 yr at 58 lakes. Sites with declining sedimentary  $\delta^{15}\text{N}$  values towards present have low MAT, high elevation and high human footprint index. Statistical results for the overall model are shown in Supplementary Table 2.

availability can stimulate  $\text{N}_2$  fixation<sup>26</sup>. Both processes could affect the signatures of sedimentary  $\delta^{15}\text{N}$  on a global scale by affecting the signature of gaseous N losses and atmospheric reactive N. For example, the  $\delta^{15}\text{N}$  of nitrate deposited on the Greenland Ice Sheet declined from 15 to 7 cal. kyr BP (ref. 27). In addition to the direct effects of increased atmospheric  $\text{CO}_2$  on local N cycling, these teleconnections could also contribute to declines in sedimentary  $\delta^{15}\text{N}$  in non-glaciated lakes such as those in eastern Africa.

The increased supply of reactive N to terrestrial ecosystems that accompanied the modern doubling of reactive-N production has been the major change in the global N cycle during the past 500 yr (ref. 3). One might expect an increase in sedimentary  $\delta^{15}\text{N}$  if demand by the terrestrial biosphere could not keep pace with the increased supply of reactive N. Yet, global increases in sedimentary  $\delta^{15}\text{N}$  during the past 500 yr in this study were not observed. One potential explanation for this might be that the modern terrestrial biosphere has been sequestering C (~4 Pg C yr<sup>-1</sup>; ref. 28), as occurred during the Pleistocene–Holocene transition. Accompanying this C sequestration, N sequestration would be approximately 36 Tg N yr<sup>-1</sup>, depending on the ecosystem C:N ratio<sup>14</sup>. Although this rate of N sequestration is ~20% of the ~185 Tg N yr<sup>-1</sup> that humans add to the biosphere<sup>3</sup>, terrestrial C sequestration and increases in atmospheric  $\text{CO}_2$  concentrations could be limiting wide-scale increases in N availability. Alternatively, the introduction of large amounts of reactive N with an atmospheric signature could be balancing the isotopic effects of fractionation due to gaseous losses. The lack of a globally consistent trend in sedimentary  $\delta^{15}\text{N}$  over the past 500 yr suggests that the ultimate outcome of these processes is being primarily controlled by local conditions. This multiplicity of responses highlights the need to understand local and regional trends in sedimentary  $\delta^{15}\text{N}$  (ref. 29), in particular to reduce uncertainties in local and global controls over the N budget and to trace the fate of deposited reactive N in terrestrial and aquatic ecosystems.

Our results add to the growing evidence that N supply and N demand must be considered jointly to understand the N status of the terrestrial biosphere. Global-scale N cycling changed long before humans significantly modified the N cycle, therefore any recent changes to the N cycle should not necessarily be interpreted as a direct effect of increased anthropogenic reactive-N production. Global biogeochemical cycles responded slowly to massive changes in planetary climate systems; it took millennia for the global N cycle to adjust to the removal of ice sheets and reorganization of the global C cycle. The dynamics of these slow processes, especially the accuracy of C accrual estimates on millennial timescales, can be studied through parameterization of C and N cycling in global vegetation and ecosystem models constrained by N isotopic data<sup>30</sup>. Overall, these data indicate an important role of net terrestrial C balance in maintaining the nutrient status of the modern biosphere.



**Figure 4 | Conceptual diagram of hypothesized drivers of isotopic signatures of sedimentary N.** a, During the early Holocene; b, modern times. Grey arrows show processes that lower the  $\delta^{15}\text{N}$  of products relative to inputs; white arrows show processes that do not alter  $\delta^{15}\text{N}$  of products. The isotopic signature of N as it is fixed from the atmosphere ( $\text{N}_2$  fix) is close to 0‰. C sequestration ( $\text{C}_{\text{seq}}$ ) and increased atmospheric  $\text{CO}_2$  concentrations could have lowered terrestrial N availability ( $-\text{N}_{\text{avail}}$ ) that would have led to fluxes of N to the atmosphere and lakes relatively depleted in  $^{15}\text{N}$ . The quantity of modern anthropogenic N addition is probably large enough to maintain similar N availability ( $=\text{N}_{\text{avail}}$ ) to ecosystems on a global scale.

## METHODS SUMMARY

**Data compilation.** Published data on lacustrine sedimentary  $\delta^{15}\text{N}$  were obtained from the corresponding author or directly from the publication. Chronologies were constructed using both  $^{210}\text{Pb}$  and  $^{14}\text{C}$  dating approaches with other fixed date horizons such as tephra. Raw radiocarbon ages were recalibrated and new age models were derived. The isotopic ratio of N ( $\delta^{15}\text{N}$ ) is expressed relative to the standard of atmospheric  $\text{N}_2$ .

**Sedimentary  $\delta^{15}\text{N}$  over time.** Sedimentary  $\delta^{15}\text{N}$  data from each site were grouped into bins to standardize sampling density (Supplementary Fig. 5). For the past 500-yr time window, sites had to meet a minimum number of 8 sample points and a duration of at least 75 yr ( $n = 58$  sites). Bin size was 10 yr. For the 15,000 yr BP time window, bin size was 100 yr and site criteria included a minimum of 10 sedimentary  $\delta^{15}\text{N}$  samples in 10 separate bins and a 1,000 yr duration ( $n = 37$ ). Nine sites met criteria for both analyses (total  $n = 86$ ). 95% confidence intervals for mean values of sedimentary  $\delta^{15}\text{N}$  were calculated by bootstrapping over 1,000 replications, with samples replaced. Breakpoint analysis was performed to identify inflection points.

**Site characteristics that influence recent trajectories.** Linear regressions between sediment age (cal. yr BP) and untransformed bulk sedimentary  $\delta^{15}\text{N}$  were performed for each site for the most recent 500 yr. Then, a stepwise forward multiple regression was performed to test the roles of site-specific variables in predicting the direction and rates of change over time among sites. Tested predictor variables included elevation (derived from the publication or a global elevation model), modelled historic and modern atmospheric N deposition levels<sup>6</sup>, mean annual temperature and mean annual precipitation, and a calculated human footprint index ranging from 0 (low impact) to 100 (high impact)<sup>16</sup>. Note that these variables reflect modern, not past, site characteristics. Variables were retained with a significance level of  $\alpha < 0.05$ .

**Full Methods** and any associated references are available in the online version of the paper.

Received 18 May 2012; accepted 15 January 2013.

- Galloway, J. N. *et al.* Nitrogen cycles: past, present, and future. *Biogeochemistry* **70**, 153–226 (2004).
- Schmitt, J. *et al.* Carbon isotope constraints on the deglacial  $\text{CO}_2$  rise from ice cores. *Science* **336**, 711–714 (2012).
- Galloway, J. N. *et al.* Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science* **320**, 889–892 (2008).
- Dentener, F. *et al.* Nitrogen and sulfur deposition on regional and global scales: a multimodel evaluation. *Glob. Biogeochem. Cycles* **20**, GB4003, <http://dx.doi.org/10.1029/2005GB002672> (2006).
- Rockström, J. *et al.* A safe operating space for humanity. *Nature* **461**, 472–475 (2009).
- Davidson, E. A. *et al.* *Excess Nitrogen in the U.S. Environment: Trends, Risks, and Solutions* Vol. 15 (Ecological Society of America, 2012).
- Bernal, S., Hedin, L. O., Likens, G. E., Gerber, S. & Buso, D. C. Complex response of the forest nitrogen cycle to climate change. *Proc. Natl Acad. Sci. USA* **109**, 3406–3411 (2012).
- Craine, J. M. *et al.* Global patterns of foliar nitrogen isotopes and their relationships with climate, mycorrhizal fungi, foliar nutrient concentrations, and nitrogen availability. *New Phytol.* **183**, 980–992 (2009).
- Bai, E. & Houlton, B. Z. Coupled isotopic and process-based modeling of gaseous nitrogen losses from tropical rain forests. *Glob. Biogeochem. Cycles* **23**, GB2011, <http://dx.doi.org/10.1029/2008GB003361> (2009).
- McLauchlan, K. K., Ferguson, C. J., Wilson, I. E., Ocheltree, T. W. & Craine, J. M. Thirteen decades of foliar isotopes indicate declining nitrogen availability in central North American grasslands. *New Phytol.* **187**, 1135–1145 (2010).
- Hietz, P. *et al.* Long-term change in the nitrogen cycle of tropical forests. *Science* **334**, 664–666 (2011).
- Perakis, S. S., Sinkhorn, E. R. & Compton, J. E.  $\delta^{15}\text{N}$  constraints on long-term nitrogen balances in temperate forests. *Oecologia* **167**, 793–807 (2011).
- Mayr, C. *et al.* Isotopic fingerprints on lacustrine organic matter from Laguna Potrok Aike (southern Patagonia, Argentina) reflect environmental changes during the last 16,000 years. *J. Paleolimnol.* **42**, 81–102 (2009).
- Houlton, B. Z. & Bai, E. Imprint of denitrifying bacteria on the global terrestrial biosphere. *Proc. Natl Acad. Sci. USA* **106**, 21713–21716 (2009).
- McLauchlan, K. K., Craine, J. M., Oswald, W. W., Leavitt, P. R. & Likens, G. E. Changes in nitrogen cycling during the past century in a northern hardwood forest. *Proc. Natl Acad. Sci. USA* **104**, 7466–7470 (2007).
- Sanderson, E. W. *et al.* The human footprint and the last of the wild. *Bioscience* **52**, 891–904 (2002).
- Prentice, I. C., Harrison, S. P. & Bartlein, P. J. Global vegetation and terrestrial carbon cycle changes after the last ice age. *New Phytol.* **189**, 988–998 (2011).
- Engstrom, D. R., Fritz, S. C., Almendinger, J. E. & Juggins, S. Chemical and biological trends during lake evolution in recently deglaciated terrain. *Nature* **408**, 161–166 (2000).
- Bunting, L., Leavitt, P. R., Weidman, R. P. & Vinebrooke, R. D. Regulation of the nitrogen biogeochemistry of mountain lakes by subsidies of terrestrial dissolved organic matter and the implications for climate studies. *Limnol. Oceanogr.* **55**, 333–345 (2010).
- Talbot, M. R. in *Tracking Environmental Change Using Lake Sediments* Vol. 2, *Physical and Geochemical Methods* (eds Last, W. M. & Smol, J. P.) 401–439 (Kluwer Academic, 2001).
- Högberg, P. Tansley Review No. 95  $^{15}\text{N}$  natural abundance in soil-plant systems. *New Phytol.* **137**, 179–203 (1997).
- Perakis, S. S. & Sinkhorn, E. R. Biogeochemistry of a temperate forest nitrogen gradient. *Ecology* **92**, 1481–1491 (2011).
- Schilt, A. *et al.* Glacial-interglacial and millennial-scale variations in the atmospheric nitrous oxide concentration during the last 800,000 years. *Quat. Sci. Rev.* **29**, 182–192 (2010).
- Petit, J. R. *et al.* Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 429–436 (1999).
- Norby, R. J., Warren, J. M., Iversen, C. M., Medlyn, B. E. & McMurtrie, R. E.  $\text{CO}_2$  enhancement of forest productivity constrained by limited nitrogen availability. *Proc. Natl Acad. Sci. USA* **107**, 19368–19373 (2010).
- Houlton, B. Z., Wang, Y. P., Vitousek, P. M. & Field, C. B. A unifying framework for dinitrogen fixation in the terrestrial biosphere. *Nature* **454**, 327–330 (2008).
- Hastings, M. G., Sigman, D. M. & Steig, E. J. Glacial/interglacial changes in the isotopes of nitrate from the Greenland Ice Sheet Project 2 (GISP2) ice core. *Glob. Biogeochem. Cycles* **19**, GB4024, <http://dx.doi.org/10.1029/2005GB002502> (2005).
- Le Quééré, C. *et al.* Trends in the sources and sinks of carbon dioxide. *Nature Geosci.* **2**, 831–836 (2009).
- Holtgrieve, G. W. *et al.* A coherent signature of anthropogenic nitrogen deposition to remote watersheds of the Northern Hemisphere. *Science* **334**, 1545–1548 (2011).
- Xu-Ri & Prentice, I. C. Terrestrial nitrogen cycle simulation with a dynamic global vegetation model. *Glob. Change Biol.* **14**, 1745–1764 (2008).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by the National Science Foundation (BCS-0955225 and EPS-0903806) and a James Martin Fellowship at the University of Oxford. We thank the many authors who contributed their data for the purpose of this analysis. We appreciate technical assistance from C. Morris, P. Long and S. McConaghy. We thank J. Marlon, S. Enders, C. Baird and S. Perakis for comments.

**Author Contributions** K.K.M. and E.S.J. designed research, K.K.M., E.S.J. and J.J.W. performed research, K.K.M., J.J.W. and J.M.C. analysed the data, and K.K.M. led the writing of the paper with substantial input from E.S.J., J.J.W. and J.M.C.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.K.M. ([mclauch@ksu.edu](mailto:mclauch@ksu.edu)).



## METHODS

**Literature search.** The search was conducted using the Web of Science database with keywords 'nitrogen isotopes' and 'lake sediments', 'delta ( $\delta^{15}\text{N}$ )' and 'lake sediments', and 'paleo' or 'palaeo' and 'nitrogen' in August 2011. Of the 627 records returned as a result of these searches, 192 met the criteria of providing well-dated stable isotope data from lacustrine sediments. All studies reported bulk sedimentary  $\delta^{15}\text{N}$  data expressed in standard delta notation relative to the isotopic composition of air<sup>31</sup>. Corresponding authors were contacted with a request for the data published in each manuscript, and in cases where an author could not be contacted and/or the original data files could not be located, the published data (sedimentary  $\delta^{15}\text{N}$  and chronology) were obtained from the manuscript. In the case of duplicate cores from a single lake, the manuscript was examined and the oldest single core was chosen. This process led to a sample size of 98 sites. If a manuscript mentioned acid pre-treatment of sediments before mass spectrometry, that site was excluded due to possible biases<sup>32</sup>. Further criteria were applied regarding temporal density of sampling to meet analytical requirements, such that the final number of sites used in this analysis was 86 lacustrine sediment records. Site information is available in Supplementary Table 1 and these references<sup>33–98</sup>.

**Establishment of chronologies.** To standardize age models for all sedimentary records, recalibration of existing radiometric ages was performed where possible. Radiocarbon ages were calibrated using the most recently produced calibration curves. SHCal04<sup>99</sup> was used for Southern Hemisphere ages 0–11 cal. kyr BP. IntCal09<sup>100</sup> was used for all Northern Hemisphere records and Southern Hemisphere ages older than 11 cal. kyr BP. For those sites where a new chronology was created, classical age–depth models were generated using CLAM<sup>101</sup> run in the open-source statistical software R<sup>102</sup>. Several age–depth models were created for each site and the choice of model used was based upon the goodness of fit, the recommendation from the original published manuscripts and the interpretation of the authors. Chronologies spanning the past 500 yr were often based partially or entirely on  $^{210}\text{Pb}$  activity using a constant rate-of-supply model. These ages were not reassessed, which might in the case of imprecise original chronologies affect the coherence of the higher-resolution 'short cores'. In order to standardize ages and to use the most accurate chronological constraint possible, all ages are given as years before present (cal. yr BP) in which '0' refers to AD 1950. Any ages more recent than this date are displayed as negative values (that is, AD 2012 = –62 cal. yr BP).

**Site characteristics and additional variables.** Latitude and longitude of each lake were acquired from the published manuscripts, as well as the surface area and elevation when provided. Two climate variables—Bio1 (mean annual temperature) and Bio12 (mean annual precipitation)—were obtained from the WorldClim data set which is presented at 1 km spatial resolution and averages AD 1950 to 2000 (<http://www.worldclim.org/>)<sup>103</sup>. Estimates of nitrogen deposited to terrestrial ecosystems from the atmosphere in 1993 and 1860 were obtained from ([http://daac.ornl.gov/CLIMATE/guides/global\\_N\\_deposition\\_maps.html](http://daac.ornl.gov/CLIMATE/guides/global_N_deposition_maps.html))<sup>104</sup>. Elevation, where not provided in manuscripts, was obtained from a global elevation model<sup>105</sup>. The human footprint metric is an index of degree of human influence on terrestrial ecosystems, calculated from data on population density, land transformation, accessibility and electrical power infrastructure. It ranges from 0 to 100 (ref. 106).

Additional geochemical data from the sediment records—N concentrations, C concentrations,  $\delta^{13}\text{C}$  and C:N—do not correlate strongly at a global scale to sedimentary  $\delta^{15}\text{N}$  values or the rate or direction of change over time (Supplementary Fig. 4). There are relationships among geochemical variables at individual sites<sup>70,107</sup>.

**Statistical analyses.** Because sedimentary  $\delta^{15}\text{N}$  is already measured relative to a global standard, minimal data manipulation was required to standardize the records. Two time windows were selected for analysis: (1) past 500 yr when atmospheric  $\text{CO}_2$  concentrations have increased by 100 p.p.m.v. (from 290 to 390 p.p.m.v.)<sup>108</sup>, and (2) the past 15,000 yr. While some sediment records were older than 15,000 yr, the data density began to degrade significantly before this time. The latter time period represents an increase in atmospheric  $\text{CO}_2$  approximately 50% of the recent increase, from 188.5 p.p.m.v. at 17,073 cal. yr BP gas age to 238.6 p.p.m.v. at 13,440 cal. yr BP gas age<sup>109</sup>.

To determine the pattern of sedimentary  $\delta^{15}\text{N}$  over time, we performed piecewise linear regressions using the 'segmented' package in R<sup>110</sup>. The existence of an inflection point, and a significant difference in slope, was tested using the Davies' test. The position, along with standard error and confidence intervals, of the inflection point was estimated (Fig. 2). Confidence intervals of 95% levels (Fig. 2) were generated by a bootstrap approach in which values for individual sites per bin window were sampled with replacement over 1,000 replications. The bootstrap approach allows assessment of the sensitivity of the results to the inclusion or exclusion of individual sites. The bootstrapping was performed in R, using the package 'Boot'<sup>111</sup>.

31. Nadelhoffer, K. J. & Fry, B. in *Stable Isotopes in Ecology and Environmental Science* (eds Laljtha, K. & Michener, R. H.) 22–44 (Blackwell Scientific, 1994).
32. Brodie, C. R. *et al.* Evidence for bias in C/N,  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  values of bulk organic matter, and on environmental interpretation, from a lake sedimentary sequence by pre-analysis acid treatment methods. *Quat. Sci. Rev.* **30**, 3076–3087 (2011).
33. Leavitt, P. R., Schindler, D. E., Paul, A. J., Hardie, A. K. & Schindler, D. W. Fossil pigment records of phytoplankton in trout-stocked alpine lakes. *Can. J. Fish. Aquat. Sci.* **51**, 2411–2423 (1994).
34. Lasca, I., McLauchlan, K. K., Myrbo, A., Leavitt, P. R. & Banerjee, S. K. Sediment-magnetic evidence for last millennium drought conditions at the prairie-forest ecotone of northern United States. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **337–338**, 99–107 (2012).
35. Laird, K., Cumming, B. & Nordin, R. A regional paleolimnological assessment of the impact of clear-cutting on lakes from the west coast of Vancouver Island, British Columbia. *Can. J. Fish. Aquat. Sci.* **58**, 479–491 (2001).
36. Laird, K. & Cumming, B. A regional paleolimnological assessment of the impact of clear-cutting on lakes from the central interior of British Columbia. *Can. J. Fish. Aquat. Sci.* **58**, 492–505 (2001).
37. Jeffers, E. S., Bonsall, M. B. & Willis, K. J. Stability in ecosystem functioning across a climatic threshold and contrasting forest regimes. *PLoS One* **6**, e16134 (2011).
38. Jeffers, E. S., Bonsall, M. B., Watson, J. E. & Willis, K. J. Climate change impacts on ecosystem functioning: evidence from an Empetrum heathland. *New Phytol.* **193**, 150–164 (2012).
39. Xu, Y. P. & Jaffe, R. Geochemical record of anthropogenic impacts on Lake Valencia, Venezuela. *Appl. Geochem.* **24**, 411–418 (2009).
40. Wooller, M., Wang, Y. M. & Axford, Y. A multiple stable isotope record of Late Quaternary limnological changes and chironomid paleoecology from northeastern Iceland. *J. Paleolimnol.* **40**, 63–77 (2008).
41. Wolfe, B. B. *et al.* Effect of varying oceanicity on early- to mid-Holocene palaeohydrology, Kola Peninsula, Russia: isotopic evidence from treeline lakes. *Holocene* **13**, 153–160 (2003).
42. Wolfe, B. B., Edwards, T. W. D. & Aravena, R. Changes in carbon and nitrogen cycling during tree-line retreat recorded in the isotopic content of lacustrine organic matter, western Taimyr Peninsula, Russia. *Holocene* **9**, 215–222 (1999).
43. Wolfe, A. P., Van Gorp, A. C. & Baron, J. S. Recent ecological and biogeochemical changes in alpine lakes of Rocky Mountain National Park (Colorado, USA): a response to anthropogenic nitrogen deposition. *Geobiology* **1**, 153–168 (2003).
44. Wolfe, A. P., Cooke, C. A. & Hobbs, W. O. Are current rates of atmospheric nitrogen deposition influencing lakes in the Eastern Canadian Arctic? *Arct. Antarct. Alp. Res.* **38**, 465–476 (2006).
45. Vreca, P. & Muri, G. Changes in accumulation of organic matter and stable carbon and nitrogen isotopes in sediments of two Slovenian mountain lakes (Lake Ledvica and Lake Planina), induced by eutrophication changes. *Limnol. Oceanogr.* **51**, 781–790 (2006).
46. Teranes, J. L. & Bernasconi, S. M. The record of nitrate utilization and productivity limitation provided by  $\delta^{15}\text{N}$  values in lake organic matter — a study of sediment trap and core sediments from Baldeggersee, Switzerland. *Limnol. Oceanogr.* **45**, 801–813 (2000).
47. Tenzer, G. E. *et al.* Sedimentary organic matter record of recent environmental changes in the St. Marys River ecosystem, Michigan-Ontario border. *Org. Geochem.* **30**, 133–146 (1999).
48. Tareq, S. M., Kitagawa, H. & Ohta, K. Lignin biomarker and isotopic records of paleovegetation and climate changes from Lake Erhai, southwest China, since 18.5 ka BP. *Quat. Int.* **229**, 47–56 (2011).
49. Talbot, M. R. & Johannessen, T. A high-resolution palaeoclimatic record for the last 27,500 years in tropical west Africa from the carbon and nitrogen isotopic composition of lacustrine organic matter. *Earth Planet. Sci. Lett.* **110**, 23–37 (1992).
50. Talbot, M. R. & Laerdal, T. The Late Pleistocene-Holocene palaeolimnology of Lake Victoria, East Africa, based upon elemental and isotopic analyses of sedimentary organic matter. *J. Paleolimnol.* **23**, 141–164 (2000).
51. Talbot, M. R., Jensen, N. B., Laerdal, T. & Filippi, M. L. Geochemical responses to a major transgression in giant African Lakes. *J. Paleolimnol.* **35**, 467–489 (2006).
52. Street-Perrott, F. A. *et al.* Towards an understanding of late Quaternary variations in the continental biogeochemical cycle of silicon: multi-isotope and sediment-flux data for Lake Rutundu, Mt Kenya, East Africa, since 38 ka BP. *J. Quat. Sci.* **23**, 375–387 (2008).
53. Selbie, D. T., Lewis, B. A., Smol, J. P. & Finney, B. P. Long-term population dynamics of the endangered Snake River sockeye salmon: evidence of past influences on stock decline and impediments to recovery. *Trans. Am. Fisheries Soc.* **136**, 800–821 (2007).
54. Selbie, D. T. Climate change modulates structural and functional lake ecosystem responses to introduced anadromous salmon. *Can. J. Fish. Aquat. Sci.* **68**, 675–692 (2011).
55. Schindler, D. E., Leavitt, P. R., Johnson, S. P. & Brock, C. S. A 500-year context for the recent surge in sockeye salmon (*Oncorhynchus nerka*) abundance in the Alagnak River, Alaska. *Can. J. Fish. Aquat. Sci.* **63**, 1439–1444 (2006).
56. Schindler, D. E., Leavitt, P. R., Brock, C. S., Johnson, S. P. & Quay, P. D. Marine-derived nutrients, commercial fisheries, and production of salmon and lake algae in Alaska. *Ecology* **86**, 3225–3231 (2005).
57. Saros, J. E., Michel, T. J., Interlandi, S. J. & Wolfe, A. P. Resource requirements of *Asterionella formosa* and *Fragilaria crotonensis* in oligotrophic alpine lakes: implications for recent phytoplankton community reorganizations. *Can. J. Fish. Aquat. Sci.* **62**, 1681–1689 (2005).
58. Ryner, M., Gasse, F., Rumes, B. & Verschuren, D. Climatic and hydrological instability in semi-arid equatorial East Africa during the late Glacial to Holocene

- transition: a multi-proxy reconstruction of aquatic ecosystem response in northern Tanzania. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **248**, 440–458 (2007).
59. Russell, J. M., McCoy, S. J., Verschuren, D., Bessems, I. & Huang, Y. Human impacts, climate change, and aquatic ecosystem response during the past 2000 yr at Lake Wandakara, Uganda. *Quat. Res.* **72**, 315–324 (2009).
  60. Routh, J., Meyers, P. A., Hjorth, T., Baskaran, M. & Hallberg, R. Sedimentary geochemical record of recent environmental changes around Lake Middle Marviken, Sweden. *J. Paleolimnol.* **37**, 529–545 (2007).
  61. Routh, J. *et al.* Sedimentary geochemical record of human-induced environmental changes in the Lake Brunnsnåven watershed, Sweden. *Limnol. Oceanogr.* **49**, 1560–1569 (2004).
  62. Routh, J., Choudhary, P., Meyers, P. A. & Kumar, B. A sediment record of recent nutrient loading and trophic state change in Lake Norrviken, Sweden. *J. Paleolimnol.* **42**, 325–341 (2009).
  63. Reuss, N. S. *et al.* Lake ecosystem responses to Holocene climate change at the subarctic tree-line in northern Sweden. *Ecosystems* **13**, 393–409 (2010).
  64. Pueyo, J. J. *et al.* Carbonate and organic matter sedimentation and isotopic signatures in Lake Chungara, Chilean Altiplano, during the last 12.3 kyr. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **307**, 339–355 (2011).
  65. Pessenda, L. C. R. *et al.* Last millennium environmental changes and climate inferences in the Southeastern Atlantic forest, Brazil. *An. Acad. Bras. Cienc.* **82**, 717–729 (2010).
  66. Parplies, J. *et al.* Late glacial environment and climate development in northeastern China derived from geochemical and isotopic investigations of the varved sediment record from Lake Sihailongwan (Jilin Province). *J. Paleolimnol.* **40**, 471–487 (2008).
  67. Olsen, J., Noe-Nygaard, N. & Wolfe, B. B. Mid- to late-Holocene climate variability and anthropogenic impacts: multi-proxy evidence from Lake Bliden, Denmark. *J. Paleolimnol.* **43**, 323–343 (2010).
  68. Muzuka, A. N. N., Ryner, M. & Holmgren, K. 12,000-year, preliminary results of the stable nitrogen and carbon isotope record from the Empakai Crater lake sediments, Northern Tanzania. *J. Afr. Earth Sci.* **40**, 293–303 (2004).
  69. McLauchlan, K. K., Craine, J. M., Oswald, W. W., Leavitt, P. R. & Likens, G. E. Changes in nitrogen cycling during the past century in a northern hardwood forest. *Proc. Natl Acad. Sci. USA* **104**, 7466–7470 (2007).
  70. Mayr, C. *et al.* Isotopic fingerprints on lacustrine organic matter from Laguna Potrok Aike (southern Patagonia, Argentina) reflect environmental changes during the last 16,000 years. *J. Paleolimnol.* **42**, 81–102 (2009).
  71. Lucke, A. & Brauer, A. Biogeochemical and micro-facial fingerprints of ecosystem response to rapid Late Glacial climatic changes in varved sediments of Meerfelder Maar (Germany). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **211**, 139–155 (2004).
  72. Lobo, I., Mozeto, A. A. & Aravena, R. Paleohydrological investigation of Infernao Lake, Moji-Guaçu River watershed, Sao Paulo, Brazil. *J. Paleolimnol.* **26**, 119–129 (2001).
  73. Abbott, M. B., Wolfe, B. B., Aravena, R., Wolfe, A. P. & Seltzer, G. O. Holocene hydrological reconstructions from stable isotopes and paleolimnology, Cordillera Real, Bolivia. *Quat. Sci. Rev.* **19**, 1801–1820 (2000).
  74. Alin, S. R. *et al.* Effects of land-use change on aquatic biodiversity: A view from the paleorecord at Lake Tanganyika, East Africa. *Geology* **30**, 1143–1146 (2002).
  75. Anderson, L., Abbott, M. B. & Finney, B. P. Holocene climate inferred from oxygen isotope ratios in lake sediments, central Brooks Range, Alaska. *Quat. Res.* **55**, 313–321 (2001).
  76. Bertrand, S. *et al.* Bulk organic geochemistry of sediments from Puyehue Lake and its watershed (Chile, 40 degrees S): implications for paleoenvironmental reconstructions. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **294**, 56–71 (2010).
  77. Brahmey, J., Clague, J. J., Menounos, B. & Edwards, T. W. D. Timing and cause of water level fluctuations in Kluane Lake, Yukon Territory, over the past 5000 years. *Quat. Res.* **70**, 213–227 (2008).
  78. Brenner, M., Whitmore, T. J., Curtis, J. H., Hodell, D. A. & Schelske, C. L. Stable isotope ( $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$ ) signatures of sedimented organic matter as indicators of historic lake trophic state. *J. Paleolimnol.* **22**, 205–221 (1999).
  79. Choudhary, P., Routh, J. & Chakrapani, G. J. An environmental record of changes in sedimentary organic matter from Lake Sattal in Kumaun Himalayas, India. *Sci. Total Environ.* **407**, 2783–2795 (2009).
  80. Choudhary, P., Routh, J., Chakrapani, G. J. & Kumar, B. Biogeochemical records of paleoenvironmental changes in Nainital Lake, Kumaun Himalayas, India. *J. Paleolimnol.* **42**, 571–586 (2009).
  81. Chu, G. Q. *et al.* A 1600 year multiproxy record of paleoclimatic change from varved sediments in Lake Xiaolongwan, northeastern China. *J. Geophys. Res.* **114**, D22108, <http://dx.doi.org/10.1029/2009JD012077> (2009).
  82. Drew, S., Flett, I., Wilson, J., Hejnis, H. & Skilbeck, C. G. The trophic history of Myall Lakes, New South Wales, Australia: interpretations using  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  of the sedimentary record. *Hydrobiologia* **608**, 35–47 (2008).
  83. Enders, S. K. *et al.* Compound-specific stable isotopes of organic compounds from lake sediments track recent environmental changes in an alpine ecosystem, Rocky Mountain National Park, Colorado. *Limnol. Oceanogr.* **53**, 1468–1478 (2008).
  84. Engstrom, D. R., Schottler, S. P., Leavitt, P. R. & Havens, K. E. A reevaluation of the cultural eutrophication of Lake Okeechobee using multiproxy sediment records. *Ecol. Appl.* **16**, 1194–1206 (2006).
  85. Enters, D. *et al.* Climate change and human impact at Sacrower See (NE Germany) during the past 13,000 years: a geochemical record. *J. Paleolimnol.* **43**, 719–737 (2010).
  86. Fagel, N. *et al.* Geochemical evidence (C, N and Pb isotopes) of recent anthropogenic impact in south-central Chile from two environmentally distinct lake sediment records. *J. Quat. Sci.* **25**, 1100–1112 (2010).
  87. Filippi, M. L. & Talbot, M. R. The paleolimnology of northern Lake Malawi over the last 25 ka based upon the elemental and stable isotopic composition of sedimentary organic matter. *Quat. Sci. Rev.* **24**, 1303–1328 (2005).
  88. Hassan, K. M., Swinehart, J. B. & Spalding, R. F. Evidence for Holocene environmental change from C/N ratios, and  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  values in Swan Lake sediments, western Sand Hills, Nebraska. *J. Paleolimnol.* **18**, 121–130 (1997).
  89. Heikkilä, M., Edwards, T. W. D., Seppä, H. & Sonninen, E. Sediment isotope tracers from Lake Saarikko, Finland, and implications for Holocene hydroclimatology. *Quat. Sci. Rev.* **29**, 2146–2160 (2010).
  90. Herczeg, A. L., Smith, A. K. & Dighton, J. C. A 120 year record of changes in nitrogen and carbon cycling in Lake Alexandrina, South Australia: C:N,  $\delta^{15}\text{N}$  and  $\delta^{13}\text{C}$  in sediments. *Appl. Geochem.* **16**, 73–84 (2001).
  91. Hobbs, W. O., Vinebrooke, R. D. & Wolfe, A. P. Biogeochemical responses of two alpine lakes to climate change and atmospheric deposition, Jasper and Banff National parks, Canadian Rocky Mountains. *Can. J. Fish. Aquat. Sci.* **68**, 1480–1494 (2011).
  92. Holmgren, S. U., Bigler, C., Ingolfsson, O. & Wolfe, A. P. The Holocene-Anthropocene transition in lakes of western Spitsbergen, Svalbard (Norwegian High Arctic): climate change and nitrogen deposition. *J. Paleolimnol.* **43**, 393–412 (2010).
  93. Holtgrieve, G. W. *et al.* A coherent signature of anthropogenic nitrogen deposition to remote watersheds of the Northern Hemisphere. *Science* **334**, 1545–1548 (2011).
  94. Hu, F. S., Finney, B. P. & Brubaker, L. B. Effects of Holocene Alnus expansion on aquatic productivity, nitrogen cycling, and soil development in southwestern Alaska. *Ecosystems* **4**, 358–368 (2001).
  95. Hyodo, F. *et al.* Changes in stable isotopes, lignin-derived phenols, and fossil pigments in sediments of Lake Biwa, Japan: implications for anthropogenic effects over the last 100 years. *Sci. Total Environ.* **403**, 139–147 (2008).
  96. Janbu, A. D., Paasche, O. & Talbot, M. R. Paleoclimate changes inferred from stable isotopes and magnetic properties of organic-rich lake sediments in Arctic Norway. *J. Paleolimnol.* **46**, 29–44 (2011).
  97. Li, L., Yu, Z., Moeller, R. E. & Bebout, G. E. Complex trajectories of aquatic and terrestrial ecosystem shifts caused by multiple human-induced environmental stress. *Geochim. Cosmochim. Acta* **72**, 4338–4351 (2008).
  98. Jeffers, E. S., Bonsall, M. B., Brooks, S. J. & Willis, K. J. Abrupt environmental changes drive shifts in tree-grass interaction outcomes. *J. Ecol.* **99**, 1063–1070 (2011).
  99. McCormac, F. G. SHCal04 Southern Hemisphere calibration 0–11.0 cal kyr BP. *Radiocarbon* **46**, 1087–1092 (2004).
  100. Reimer, P. J. *et al.* IntCal09 and Marine09 radiocarbon age calibration curves, 0–50,000 years cal BP. *Radiocarbon* **51**, 1111–1150 (2009).
  101. Blaauw, M. Methods and code for “classical” age-modelling of radiocarbon sequences. *Quat. Geochronol.* **5**, 512–518 (2010).
  102. Team, R. D. C. R. *A Language and Environment for Statistical Computing. Reference Index Version 2.15.0* (R Foundation for Statistical Computing, 2005).
  103. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
  104. Dentener, F. *et al.* Nitrogen and sulfur deposition on regional and global scales: a multimodel evaluation. *Glob. Biogeochem. Cycles* **20**, GB4003, <http://dx.doi.org/10.1029/2005GB002672> (2006).
  105. Farr, T. G. *et al.* The shuttle radar topography mission. *Rev. Geophys.* **45**, RG2004 (2007).
  106. Sanderson, E. W. *et al.* The human footprint and the last of the wild. *Bioscience* **52**, 891–904 (2002).
  107. Woodward, C. A., Potito, A. P. & Beilman, D. W. Carbon and nitrogen stable isotope ratios in surface sediments from lakes of western Ireland: implications for inferring past lake productivity and nitrogen loading. *J. Paleolimnol.* **47**, 167–184 (2012).
  108. Keeling, C. D., Piper, S. C., Whorf, T. P. & Keeling, R. F. Evolution of natural and anthropogenic fluxes of atmospheric CO<sub>2</sub> from 1957 to 2003. *Tellus B* **63**, 1–22 (2011).
  109. Petit, J. R. *et al.* Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 429–436 (1999).
  110. Vito, M. & Muggeo, R. segmented: an R package to fit regression models with broken-line relationships. *R News* **8**, 20–25 (2008).
  111. Canty, A. & Ripley, B. Resampling methods in R: the boot package. *R News* **2**, 2–7 (2002).



# Melt-rich channel observed at the lithosphere–asthenosphere boundary

S. Naif<sup>1</sup>, K. Key<sup>1</sup>, S. Constable<sup>1</sup> & R. L. Evans<sup>2</sup>

The lithosphere–asthenosphere boundary (LAB) separates rigid oceanic plates from the underlying warm ductile asthenosphere. Although a viscosity decrease beneath this boundary is essential for plate tectonics, a consensus on its origin remains elusive. Seismic studies identify a prominent velocity discontinuity at depths thought to coincide with the LAB but disagree on its cause<sup>1–5</sup>, generally invoking either partial melting<sup>6</sup> or a mantle dehydration boundary<sup>7</sup> as explanations. Here we use sea-floor magnetotelluric data to image the electrical conductivity of the LAB beneath the edge of the Cocos plate at the Middle America trench offshore of Nicaragua. Underneath the resistive oceanic lithosphere, the magnetotelluric data reveal a high-conductivity layer confined to depths of 45 to 70 kilometres. Because partial melts are stable at these depths in a warm damp mantle<sup>8</sup>, we interpret the conductor to be a partially molten layer capped by an impermeable frozen lid that is the base of the lithosphere. A conductivity anisotropy parallel to plate motion indicates that this melt has been sheared into flow-aligned tube-like structures<sup>9</sup>. We infer that the LAB beneath young plates consists of a thin, partially molten, channel of low viscosity that acts to decouple the overlying brittle lithosphere from the deeper convecting mantle. Because this boundary layer has the potential to behave as a lubricant to plate motion, its proximity to the trench may have implications for subduction dynamics.

The passive magnetotelluric method uses naturally occurring magnetic fields and induced electric fields to probe mantle electrical conductivity structure. Because pure melts are at least two orders of magnitude more conductive than a typical peridotite, mantle conductivity can be greatly increased by small amounts of partial melt<sup>10</sup>, whereas conductivity depends to a much lesser extent on mantle temperature and hydration state<sup>11</sup>. This large conductivity contrast makes magnetotelluric soundings highly sensitive to the presence of mantle partial melts.

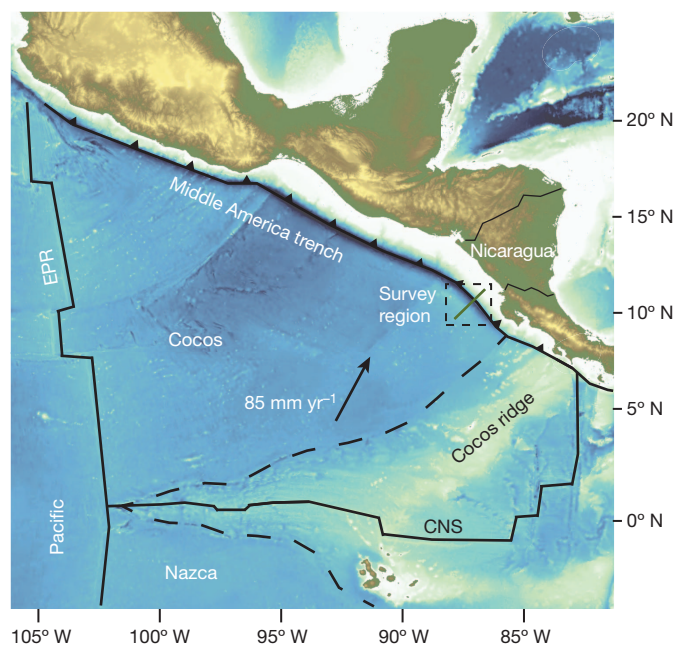
We deployed 50 broadband electromagnetic stations across a single 280-km-long profile that extended from the Cocos plate abyssal plain, across the trench, and onto the continental shelf, amounting to the largest subduction zone electromagnetic deployment to date (Fig. 1). Data from high-frequency controlled-source electromagnetic profiling revealed that the pervasive extensional bending faults on the trench outer rise are porous channels that allow for deep crustal hydration and serpentinization of the uppermost mantle<sup>12</sup>. The lower-frequency magnetotelluric data considered here target deeper conductivity structure in the lithosphere and asthenosphere. We extended the instrument array far onto the abyssal plain to image what we expected to be unremarkable conductivity beneath the oceanic plate, which would provide a baseline for comparisons with conductivity features found beneath the trench axis and continental margin.

We converted the observed magnetotelluric responses into a two-dimensional conductivity model using a nonlinear regularized inversion that solved for the triaxially anisotropic conductivity tensor of 10,000 mesh cells<sup>13,14</sup>. Figure 2a shows the horizontal resistivity (reciprocal of

conductivity) component aligned in the direction of plate motion. All three tensor components are shown in Supplementary Fig. 3.

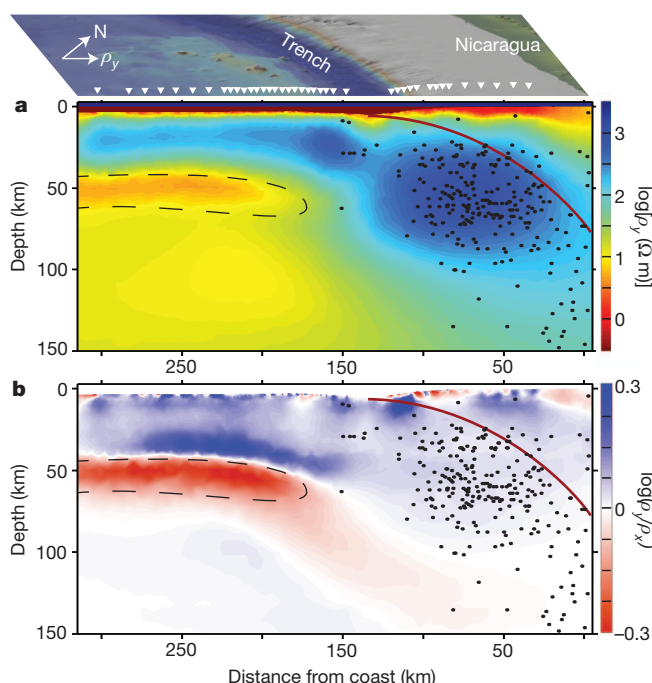
Landward of the trench, the magnetotelluric data reveal a resistive subducting slab and mantle, with a notable correlation between the location of earthquakes and high resistivity (Fig. 2a). The widespread distribution of earthquakes indicates that the brittle slab is rupturing over a wide depth range (rather than the rupturing being concentrated along the plate interface), consistent with high resistivity indicating a relatively cold, fluid-free mantle.

The conductivity becomes laterally uniform westward of the trench. A veneer of low-resistivity sediments and extrusive volcanics overlie a highly resistive lithosphere, typical of oceanic plate structure<sup>15</sup>. Below this, we find an anomalously conductive, horizontally extensive layer of 4–6  $\Omega$  m at 45–70 km depth. This conductor extends at least to the western edge of the profile where our magnetotelluric array ends, while its eastern edge lies beneath the trench outer rise wall. The anomalous layer is 1.5–2 times more conductive in the direction parallel to plate motion than parallel to the trench axis (Fig. 2b.) The



**Figure 1 | Regional tectonic map and location of the magnetotelluric survey.** The 280-km magnetotelluric profile (boxed) crosses the Middle America trench offshore of Nicaragua, where 23–24-Myr-old Cocos plate subducts beneath the Caribbean plate at a rate of 85 mm yr<sup>−1</sup>. The black dashed line separates the portions of the Cocos plate produced by either the East Pacific Rise (EPR) or the Cocos-Nazca spreading centre (CNS). The 50 sea-floor magnetotelluric stations were deployed at 10- and 4-km intervals in water depths ranging from 60 to 5,150 m and recorded magnetotelluric signals for 19 days during the month-long survey.

<sup>1</sup>Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, USA. <sup>2</sup>Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.



**Figure 2 | Resistivity model obtained from anisotropic inversion of the sea-floor magnetotelluric data.** At the top is the surface view; arrows show the direction of north and  $\rho_y$  (see below), and inverted triangles denote seafloor magnetotelluric station locations. **a**, The electrical resistivity in the direction parallel to plate motion ( $\rho_y$ ). The colour scale gives  $\log[\rho_y (\Omega \text{ m})]$ , with blue and red colours corresponding to resistive and conductive (less resistive) features, respectively. The dark red line is a model of the top of the subducting slab<sup>30</sup>. Earthquake hypocentres from up to 50 km off-axis are shown as black circles (from the USGS/NEIC catalogue). The region enclosed by the dashed black line is where the model is at least 1.5 times more conductive in the direction parallel to plate motion. **b**, Resistivity ratio for the plate-motion-parallel ( $\rho_y$ ) to trench-axis-parallel ( $\rho_x$ ) model components. The colour scale gives  $\log(\rho_y/\rho_x)$ , and the plot shows the strong anisotropy of the conductive layer at 45–70 km depth (red regions >150 km offshore). Although the lithosphere above shows a strong anisotropy, we warn that this is not well constrained, because the magnetotelluric method is primarily sensitive to conductive rather than resistive features<sup>20</sup>. The deeper mantle beneath the conductive layer is isotropic, suggesting it is not being sheared.

deeper asthenosphere is isotropic, with resistivity 10–20  $\Omega \text{ m}$ . Model sensitivity studies indicate that the conductance (the product of conductivity and thickness) of the anisotropic conductive layer and the conductivity of the underlying isotropic asthenosphere are well constrained by the data (Supplementary Fig. 4). Because the magnetotelluric data are primarily sensitive to the layer conductance, the response of the observed anomalous layer, which is 25 km thick, is nearly equivalent to that from a 12.5 km layer with a conductivity twice as large; however, we favour the thicker layer found by the inversion because its higher resistivity requires less partial melt to explain it.

The anomalous layer is too conductive to be consistent with localized temperature variations, but could be explained by either a hydrated mantle or a small degree of partial melt. Mantle olivine with a water content of 800 parts per million (p.p.m.) has a similar conductivity to the observed layer<sup>11</sup> (Fig. 3a), but such a high water content would lead to significant partial melting over this depth range<sup>16</sup> (Fig. 3c). This leads us to conclude that the conductive layer must contain partial melt, arising from a mantle with a much lower degree of hydration.

Assuming a 23-Myr-old oceanic lithosphere with a mantle potential temperature of 1,420 °C, a partially molten peridotite matrix is stable beneath our profile (Fig. 3c)<sup>16,17</sup>. The top of the melting region at 45 km depth agrees with the intersection of the 23-Myr-old oceanic plate geotherm and a damp solidus containing  $275 \pm 85$  p.p.m.  $\text{H}_2\text{O}$ , signifying the minimum water content required for the stability of melt

to coincide with our observed high-conductivity layer. This intersection may culminate in a freezing front where melt solidifies owing to the colder temperatures above. We infer that this freezing front forms a permeability barrier that traps buoyant melt beneath, rather than allowing it to percolate to shallower depths. This barrier is further reinforced by the higher viscosity of the uppermost mantle above; this higher viscosity occurs because the uppermost mantle has been depleted of volatiles during upwelling and melting at the mid-ocean ridge<sup>18</sup>. This upper boundary may also represent a sharp increase in the solubility of water in mantle minerals, resulting in a corresponding sharp decrease in the stability of partial melt<sup>19</sup>.

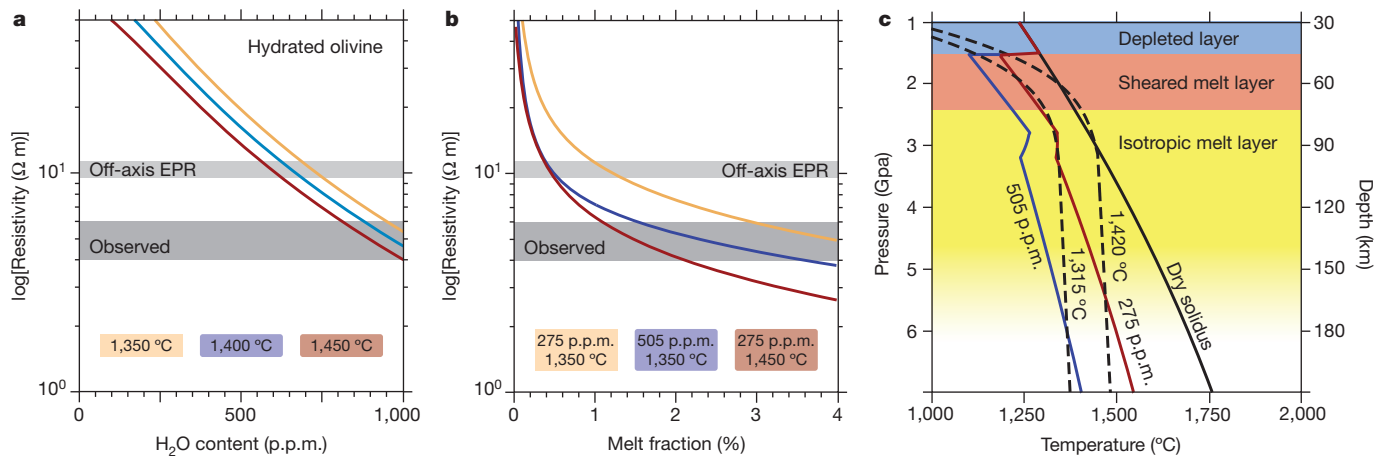
Estimates derived from samples of mid-ocean ridge basalts give mantle  $\text{H}_2\text{O}$  contents of 50–200 p.p.m., which are too low for the 275 p.p.m.  $\text{H}_2\text{O}$  needed to sustain partial melts at 45 km depth. However, uncertainties in partition coefficient measurements may accommodate this discrepancy; the resulting statistical lower bound of 190 p.p.m.  $\text{H}_2\text{O}$  is within the range of estimates from mid-ocean-ridge basalts. We do not consider the effects that other volatiles (specifically,  $\text{CO}_2$ ) have on the solidus, which could further reduce the concentration of water required to induce silicate melts at 45–70-km depth<sup>8</sup>.

The electrical anisotropy indicates that the melt is made up of a network of tubes or elongated spheroids aligned in the direction of plate motion<sup>9</sup>. Hence, we estimate the amount of partial melt present by using a parallel mixing model appropriate for anisotropic conductivity, together with results of recent laboratory studies of hydrated basaltic melt conductivity<sup>10</sup>, yielding melt fractions of 1.0–2.1% for a mantle containing 275 p.p.m.  $\text{H}_2\text{O}$  (Fig. 3b). Further reductions in melt estimates are possible, but require water concentrations that may be unrealistically high. For instance, 0.3% melt is attained for a mantle with 720 p.p.m.  $\text{H}_2\text{O}$ . Likewise, carbonatite melt is significantly more conductive and thus yields even smaller fractions, but is an unlikely explanation because it is unstable at these depths<sup>8</sup>. Bulk conductivity in the deeper isotropic asthenosphere is best described by the Hashin–Shtrikman upper bound, which predicts well-connected melt fractions of 0.2–1.1% for the observed 10–20  $\Omega \text{ m}$  mantle.

Deep off-axis melt emplacement can occur during lithosphere formation at the ridge or result from accumulation of a small degree of intraplate melting<sup>8</sup>. An earlier magnetotelluric experiment at the southern East Pacific Rise spreading ridge observed an off-axis conductive asthenosphere at 60–120 km depth beneath a resistive lithosphere<sup>20,21</sup>. We interpret that conductive layer as requiring partial melt, because its conductivity and anisotropy are similar to observations beneath our profile (Fig. 3b). That portion of the ridge generates the Nazca plate; if we assume equivalent ridge processes are occurring for the Cocos plate, where our profile was measured, then most of the hydrous melt has been emplaced at or near the ridge axis. Additional melt may accumulate through deeper intraplate melting—for example, from small-scale convection<sup>22</sup>. Grain boundary migration of deeper partial melt<sup>23</sup> carries with it increased water content as a result of the preferential partitioning of water into silicate melts. As enriched hydrated melt rises, it collects beneath the colder, less permeable lithosphere<sup>24</sup> and shears into a network of interconnected horizontally aligned melt bands<sup>25</sup>, possibly driven by large-scale asthenospheric flow<sup>26</sup>. Stress gradients perpendicular to the shear direction have been shown to reduce the melt connectivity in the direction along the gradient<sup>9</sup>, offering a possible mechanism for our observed anisotropy. The depth extent over which the asthenosphere is being sheared can be inferred from the anisotropy of the melt layer along with its increased conductivity, which the magnetotelluric data constrain to a maximum thickness of 30 km.

The existence of a horizontally extensive melt layer that is being sheared over a confined depth interval indicates the LAB is a thin, low-viscosity channel. Theoretical studies suggest that even small melt fractions (<1%) can lower viscosity by up to two orders of magnitude, effectively decoupling the lithosphere from the asthenosphere<sup>27</sup>. The depth extent over which this decoupling occurs as well as its proximity





**Figure 3 | High asthenosphere conductivity explained by a thin partially molten layer.** Evidence for the stability of melt. **a**, Hydrous olivine resistivity is estimated as a function of H<sub>2</sub>O content for different temperatures<sup>11</sup>. The dark grey region represents the observed resistivity (4–6 Ω m) of the anomalous conductive layer, while the lighter grey region represents a 10 Ω m conductive layer observed off-axis at the southern EPR<sup>20</sup>. Olivine would be required to contain at least 800 p.p.m. H<sub>2</sub>O to account for the observed resistivity if no melt is present. **b**, Bulk resistivity of partial melt shown as a function of melt fraction for different temperatures and bulk mantle water contents<sup>10</sup> with an assumed

to the trench axis have potentially new implications for plate dynamics, as previous studies have focused on the effects of melt at the ridge axis and mantle wedge. A pervasive feature that exists beneath oceanic plates at large distances from spreading centres requires the stability of a partially molten layer, as dictated by the solidus of wet peridotite and a warm mantle geotherm. A previous magnetotelluric study of 140–150-Myr-old lithosphere in the Pacific Ocean basin did not find a conductive melt layer<sup>28</sup>. This is consistent with a thicker cold plate containing too little water to sustain hydrous melts at the greater LAB depth.

In our profile, the conductive layer ends near the trench, whereas the layer anisotropy persists in a trajectory that follows the subducting slab, albeit with a decreasing anisotropic factor. This signifies that as the plate subducts, some melt is likely to remain in place beneath the shallower LAB at the trench outer rise owing to its buoyancy. This offers an explanation in which the observed melt-rich LAB is caused by the concentration of low-fraction melts at the trench over time as the plate continues to be subducted; if this mechanism is responsible for locally enriching the asthenosphere with melt, we predict that melt would be most concentrated near the trench and decrease towards younger regions of the plate.

## METHODS SUMMARY

We processed the magnetotelluric data using a standard robust multi-station impedance estimation approach at periods of 22–32,000 s (ref. 29). Supplementary Fig. 1 presents the data and two-dimensional inversion model fits. The data subset suitable for two-dimensional modelling was evaluated by inspecting impedance polarization diagrams (Supplementary Fig. 2). Sites S31–S34, located on the margin slope near the base of the trench, display strong three-dimensional effects and were omitted entirely. The rest of the data exhibit a profile-wide rotation in strike angle at periods greater than 1,024 s that is incompatible with two-dimensional modelling and were also omitted.

The two-dimensional triaxially anisotropic inversion model was obtained using a standard inversion approach implemented using a new parallel adaptive finite element algorithm<sup>13,14</sup>. In addition to penalizing spatial variations in the model resistivity, the inversion included a penalty to minimize the amount of anisotropy produced in the inverted model. The data uncertainty was subject to a 10% error floor, with the inversion finding a smooth model with a root-mean-square (r.m.s.) misfit of 1.0 for the combined TE and TM mode data. Supplementary Fig. 3 shows the resistivity obtained for each coordinate axis.

Three tests were performed to investigate the sensitivity of the model to the interpreted features (Supplementary Fig. 4). All attempts significantly increased

mineral/melt partition coefficient of 0.006. **c**, Solid lines depict the solidus of dry and wet peridotite for various mantle H<sub>2</sub>O contents<sup>16</sup>. The dashed lines show the geotherms derived from a plate cooling model of 23-Myr-old oceanic lithosphere<sup>17</sup>. Melt is stable at depths greater than 45 km for peridotite with 275 ± 85 p.p.m. H<sub>2</sub>O and a 1,420 °C mantle potential temperature. A geotherm with a reduced mantle potential temperature of 1,315 °C requires 505 ± 155 p.p.m. H<sub>2</sub>O. The solidus of wet peridotite is calculated with the cryoscopic approximation assuming an oxide molar mass<sup>16</sup>.

the r.m.s. misfit, confirming the strong data sensitivity to the interpreted features of the converged model.

The solidus of wet peridotite was derived using the cryoscopic approximation with a molar enthalpy of fusion formulation (as described in ref. 16). The mole fraction of water in the melt was calculated with an oxide molar unit (59 g mol<sup>-1</sup>). The solidi are compared to geotherms calculated from a plate cooling model that has a 95-km-thick plate with a 1,420 °C or 1,315 °C mantle potential temperature for 23-Myr-old oceanic plate<sup>17</sup>.

Received 16 October 2012; accepted 28 January 2013.

- Rychert, C. A. & Shearer, P. M. A global view of the lithosphere-asthenosphere boundary. *Science* **324**, 495–498 (2009).
- Schmerr, N. The Gutenberg discontinuity: melt at the lithosphere-asthenosphere boundary. *Science* **335**, 1480–1483 (2012).
- Bagley, B. & Revenaugh, J. Upper mantle seismic shear discontinuities of the Pacific. *J. Geophys. Res.* **113**, B12301, <http://dx.doi.org/10.1029/2008JB005692> (2008).
- Nettles, M. & Dziewoński, A. M. Radially anisotropic shear velocity structure of the upper mantle globally and beneath North America. *J. Geophys. Res.* **113**, B02303, <http://dx.doi.org/10.1029/2006JB004819> (2008).
- Kawakatsu, H. *et al.* Seismic evidence for sharp lithosphere-asthenosphere boundaries of oceanic plates. *Science* **324**, 499–502 (2009).
- Anderson, D. L. & Sammis, C. Partial melting in the upper mantle. *Phys. Earth Planet. Inter.* **3**, 41–50 (1970).
- Karato, S. & Jung, H. Water, partial melting and the origin of the seismic low velocity and high attenuation zone in the upper mantle. *Earth Planet. Sci. Lett.* **157**, 193–207 (1998).
- Hirschmann, M. M. Partial melt in the oceanic low velocity zone. *Phys. Earth Planet. Inter.* **179**, 60–71 (2010).
- Caricchi, L., Gaillard, F., Mecklenburgh, J. & Le Trong, E. Experimental determination of electrical conductivity during deformation of melt-bearing olivine aggregates: implications for electrical anisotropy in the oceanic low velocity zone. *Earth Planet. Sci. Lett.* **302**, 81–94 (2011).
- Ni, H., Keppler, H. & Behrens, H. Electrical conductivity of hydrous basaltic melts: implications for partial melting in the upper mantle. *Contrib. Mineral. Petrol.* **162**, 637–650 (2011).
- Poe, B. T., Romano, C., Nestola, F. & Smyth, J. R. Electrical conductivity anisotropy of dry and hydrous olivine at 8 GPa. *Phys. Earth Planet. Inter.* **181**, 103–111 (2010).
- Key, K., Constable, S., Matsuno, T., Evans, R. L. & Myer, D. Electromagnetic detection of plate hydration due to bending faults at the Middle America Trench. *Earth Planet. Sci. Lett.* **351–352**, 45–53 (2012).
- deGroot-Hedlin, C. & Constable, S. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics* **55**, 1613–1624 (1990).
- Key, K. & Oval, J. A parallel goal-oriented adaptive finite element method for 2.5-D electromagnetic modelling. *Geophys. J. Int.* **186**, 137–154 (2011).
- Cox, C. S., Constable, S. C., Chave, A. D. & Webb, S. C. Controlled-source electromagnetic sounding of the oceanic lithosphere. *Nature* **320**, 52–54 (1986).

16. Hirschmann, M. M., Tenner, T., Aubaud, C. & Withers, A. C. Dehydration melting of nominally anhydrous mantle: the primacy of partitioning. *Phys. Earth Planet. Inter.* **176**, 54–68 (2009).
17. Stein, C. A. & Stein, S. A model for the global variation in oceanic depth and heat flow with lithospheric age. *Nature* **359**, 123–129 (1992).
18. Hirth, G. & Kohlstedt, D. L. Water in the oceanic upper mantle: implications for rheology, melt extraction and the evolution of the lithosphere. *Earth Planet. Sci. Lett.* **144**, 93–108 (1996).
19. Mierdel, K., Keppler, H., Smyth, J. R. & Langenhorst, F. Water solubility in aluminous orthopyroxene and the origin of Earth's asthenosphere. *Science* **315**, 364–368 (2007).
20. Evans, R. L. *et al.* Geophysical evidence from the MELT area for compositional controls on oceanic plates. *Nature* **437**, 249–252 (2005).
21. Baba, K., Chave, A. D., Evans, R. L., Hirth, G. & Mackie, R. L. Mantle dynamics beneath the East Pacific Rise at 17°S: insights from the Mantle Electromagnetic and Tomography (MELT) experiment. *J. Geophys. Res.* **111**, B02101, <http://dx.doi.org/10.1029/2004JB003598> (2006).
22. Ballmer, M. D., van Hunen, J., Ito, G., Tackley, P. J. & Bianco, T. A. Non-hotspot volcano chains originating from small-scale sublithospheric convection. *Geophys. Res. Lett.* **34**, L23310, <http://dx.doi.org/10.1029/2007GL031636> (2007).
23. Zhu, W., Gaetani, G. A., Fosseis, F., Montesi, L. G. J. & De Carlo, F. Microtomography of partially molten rocks: three-dimensional melt distribution in mantle peridotite. *Science* **332**, 88–91 (2011).
24. Katz, R. F. Magma dynamics with the enthalpy method: benchmark solutions and magmatic focusing at mid-ocean ridges. *J. Petrol.* **49**, 2099–2121 (2008).
25. Kohlstedt, D. L. & Holtzman, B. K. Shearing melt out of the Earth: an experimentalist's perspective on the influence of deformation on melt extraction. *Annu. Rev. Earth Planet. Sci.* **37**, 561–593 (2009).
26. Höink, T., Jellinek, A. M. & Lenardic, A. Viscous coupling at the lithosphere–asthenosphere boundary. *Geochem. Geophys. Geosyst.* **12**, Q0AK02 (2011).
27. Takei, Y. & Holtzman, B. K. Viscous constitutive relations of solid-liquid composites in terms of grain boundary contiguity: 1. Grain boundary diffusion control model. *J. Geophys. Res.* **114**, B06205, <http://dx.doi.org/10.1029/2008JB005850> (2009).
28. Matsuno, T. *et al.* Upper mantle electrical resistivity structure beneath the central Mariana subduction system. *Geochem. Geophys. Geosyst.* **11**, Q09003 (2010).
29. Egbert, G. D. Robust multiple-station magnetotelluric data processing. *Geophys. J. Int.* **130**, 475–496 (1997).
30. Hayes, G. P., Wald, D. J. & Johnson, R. L. Slab1.0: A three-dimensional model of global subduction zone geometries. *J. Geophys. Res.* **117**, B01302, <http://dx.doi.org/10.1029/2011JB008524> (2012).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank the captain (M. Stein) and crew of R/V *Melville* and the governments of Nicaragua and Costa Rica for permission to work in their exclusive economic zones. The following are thanked for their participation in the research cruise: C. Armerding, C. Berger, E. Carruthers, B. Cohen, J. Elsenbeck, T. Matsuno, D. Myer, A. Orange, J. Perez, K. Shadle, J. Souders, K. Weitemeyer, B. Wheelock and S. Zipper; J. Lemire and A. Jacobs are thanked for their efforts with cruise planning, mobilization and demobilization. We thank B. Wheelock and D. Hasterok for discussions. This work was supported by the National Science Foundation (grants OCE-08411141 and OCE-0840894) and the Seafloor Electromagnetic Methods Consortium at Scripps Institution of Oceanography.

**Author Contributions** K.K., R.L.E. and S.C. conceived the survey. K.K. and S.C. collected the data. S.N. and S.C. processed the data. S.N. analysed and inverted the data. S.N. and K.K. developed the interpretation and wrote the manuscript. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.N. ([снаиф@ucsd.edu](mailto:снаиф@ucsd.edu)).



# The genomic signature of dog domestication reveals adaptation to a starch-rich diet

Erik Axelsson<sup>1</sup>, Abhirami Ratnakumar<sup>1</sup>, Maja-Louise Arendt<sup>1</sup>, Khurram Maqbool<sup>1</sup>, Matthew T. Webster<sup>1</sup>, Michele Perloski<sup>2</sup>, Olof Liberg<sup>3</sup>, Jon M. Arnemo<sup>4,5</sup>, Åke Hedhammar<sup>6</sup> & Kerstin Lindblad-Toh<sup>1,2</sup>

**The domestication of dogs was an important episode in the development of human civilization. The precise timing and location of this event is debated<sup>1–5</sup> and little is known about the genetic changes that accompanied the transformation of ancient wolves into domestic dogs. Here we conduct whole-genome resequencing of dogs and wolves to identify 3.8 million genetic variants used to identify 36 genomic regions that probably represent targets for selection during dog domestication. Nineteen of these regions contain genes important in brain function, eight of which belong to nervous system development pathways and potentially underlie behavioural changes central to dog domestication<sup>6</sup>. Ten genes with key roles in starch digestion and fat metabolism also show signals of selection. We identify candidate mutations in key genes and provide functional support for an increased starch digestion in dogs relative to wolves. Our results indicate that novel adaptations allowing the early ancestors of modern dogs to thrive on a diet rich in starch, relative to the carnivorous diet of wolves, constituted a crucial step in the early domestication of dogs.**

Domestic animals are crucial to modern human society, and it is likely that the first animal to be domesticated was the dog. Claims of early, fossilised dog remains include a 33,000-year-old doglike canid from the Altai Mountains in Siberia<sup>1</sup>, whereas fossils dating from 12,000–11,000 years BP found buried together with humans in Israel<sup>2</sup> could represent the earliest verified dog remains. Patterns of genomic variation indicate that dog domestication started at least 10,000 years BP<sup>3,4</sup> in southern East Asia<sup>3</sup> or the Middle East<sup>5</sup>. Dog domestication may however have been more complex, involving multiple source populations and/or backcrossing with wolves.

It is unclear why and how dogs were domesticated. Humans may have captured wolf pups for use in guarding or hunting, resulting in selection for traits of importance for these new roles. Alternatively, as humans changed from a nomadic to sedentary lifestyle during the dawn of the agricultural revolution, wolves may themselves have been attracted to dumps near early human settlements to scavenge<sup>6</sup>. Natural selection for traits allowing for efficient use of this new resource may have led to the evolution of a variety of scavenger wolves that constituted the ancestors of modern dogs. Regardless of how dog domestication started, several characteristics separating modern dogs from wolves, including reduced aggressiveness and altered social cognition capabilities<sup>7</sup>, suggest that behavioural changes were early targets of this process<sup>6</sup>. Dogs also differ morphologically from wolves, showing reduced skull, teeth and brain sizes<sup>6</sup>. Artificial selection for tameness in silver foxes indicates that selection on genetic variation in developmental genes may underlie both behavioural and morphological changes, potentially representing an important mechanism throughout animal domestication<sup>7,8</sup>.

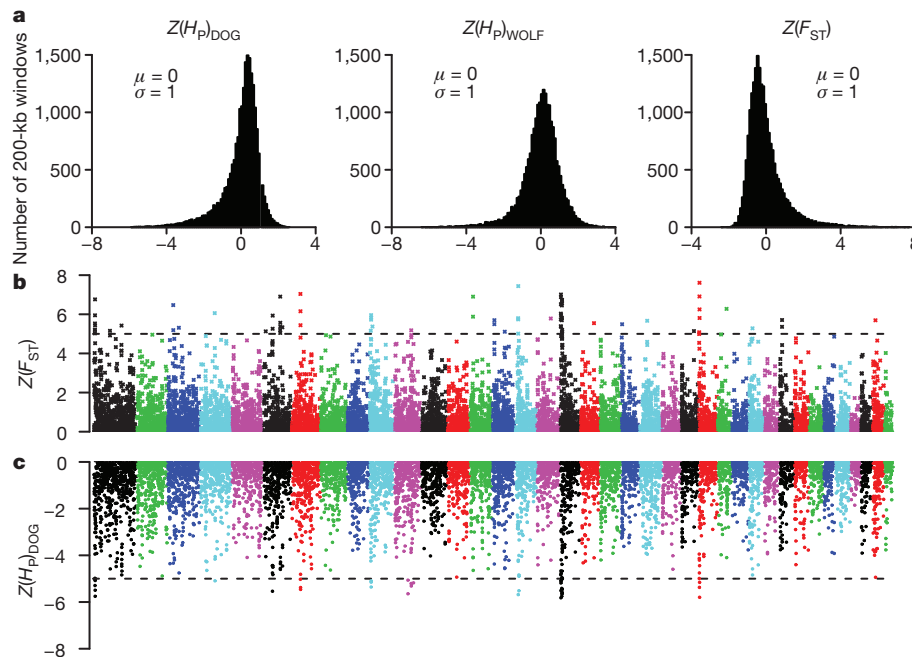
At present, only a handful of genes separating wild from domestic forms have been identified in any domestic animals, including coat

colour variants in *MC1R* in pig<sup>9</sup> and a mutation in *TSHR* likely to affect seasonal reproduction in chicken<sup>10</sup>, but to our knowledge in dogs no genome-wide sequence-based searches have been performed until now. To identify genomic regions under selection during dog domestication we performed pooled whole-genome resequencing of dogs and wolves followed by functional characterization of candidate genes.

Uniquely placed sequence reads from pooled DNA representing 12 wolves of worldwide distribution and 60 dogs from 14 diverse breeds (Supplementary Table 1) covered 91.6% and 94.6%, respectively, of the 2,385 megabases (Mb) of autosomal sequence in the CanFam 2.0 genome assembly<sup>11</sup>. The aligned coverage depth was 29.8× for all dog pools combined and 6.2× for the single wolf pool (Supplementary Table 1 and Supplementary Fig. 1). We identified 3,786,655 putative single nucleotide polymorphisms (SNPs) in the combined dog and wolf data, 1,770,909 (46.8%) of which were only segregating in the dog pools, whereas 140,818 (3.7%) were private to wolves (Supplementary Table 2). Similarly we detected 506,148 short indels and 26,619 copy-number variations (CNVs) (Supplementary Files 1 and 2). We were able to experimentally validate 113 out of 114 tested SNPs (Supplementary Table 3 and Supplementary Discussion, section 1).

To detect signals of strong recent selection we searched the dog genome for regions with reduced pooled heterozygosity ( $H_p$ )<sup>10</sup> and/or increased genetic distance to wolf ( $F_{ST}$ ). As evident from the skewed distribution of heterozygosity scores in dog relative to wolf (Fig. 1a and Supplementary Fig. 2), a major challenge to this approach is to separate true signals of selection from those caused by random fixation of large genomic regions during the formation of dog breeds<sup>11</sup>. We alleviate this problem by combining sequence data from all dog pools before selection analyses and require that detected signals span at least 200 kilobases (kb; Methods and Supplementary Discussion, sections 2 and 3). Given the complex and partly unknown demographic history of dogs, it is furthermore difficult to assign strict thresholds that distinguish selection and drift. We propose that the best way to validate regions detected here is to study genetic data from additional individuals and provide evidence for functional change associated with putatively selected regions. Eventually, indications that similar pathways changed during independent domestication events may provide conclusive evidence for selection. Here we Z-transform the autosomal  $H_p$  ( $Z(H_p)$ ) and  $F_{ST}$  ( $Z(F_{ST})$ ) distributions (see Supplementary Discussion, section 4 for an analysis of the X chromosome) and focus our description of putatively selected regions to those that fall at least five standard deviations away from the mean ( $Z(H_p) < -5$  and  $Z(F_{ST}) > 5$ ), as these represent the extreme ends of the distributions. By applying these thresholds we identified 14 regions in the dog genome with extremely low levels of heterozygosity (average length = 400 kb, average  $H_p$  = 0.036 (range 0.015–0.056), average autosomal  $H_p$  = 0.331) (Fig. 1c and Supplementary Table 4) and 35 regions with strongly elevated  $F_{ST}$  values (average length = 340 kb, average

<sup>1</sup>Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, 75237 Uppsala, Sweden. <sup>2</sup>Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts 02139, USA. <sup>3</sup>Grimsö Wildlife Research Station, Department of Ecology, Swedish University of Agricultural Sciences, 73091 Riddarhyttan, Sweden. <sup>4</sup>Department of Forestry and Wildlife Management, Faculty of Applied Ecology and Agricultural Sciences, Hedmark University College, Campus Evenstad, NO-2418 Elverum, Norway. <sup>5</sup>Department of Wildlife, Fish and Environmental Studies, Faculty of Forest Sciences, Swedish University of Agricultural Sciences, 901 83 Umeå, Sweden. <sup>6</sup>Science for Life Laboratory, Department of Clinical Sciences, Swedish University of Agricultural Sciences, 75651 Uppsala, Sweden.



**Figure 1 | Selection analyses identified 36 candidate domestication regions.** **a**, Distribution of  $Z$ -transformed average pooled heterozygosity in dog ( $Z(H_p)_{\text{DOG}}$ ) and wolf ( $Z(H_p)_{\text{WOLF}}$ ) respectively, as well as average fixation index ( $Z(F_{\text{ST}})$ ), for autosomal 200 kb windows ( $\sigma$ , standard deviation;  $\mu$ , average). **b**, The positive end of the  $Z(F_{\text{ST}})$  distribution plotted along dog

autosomes 1–38 (chromosomes are separated by colour). A dashed horizontal line indicates the cut-off ( $Z > 5$ ) used for extracting outliers. **c**, The negative end of the  $Z(H_p)$  distribution plotted along dog autosomes 1–38. A dashed horizontal line indicates the cut-off ( $Z < -5$ ) used for extracting outliers.

$F_{\text{ST}} = 0.734$  (range 0.654–0.903), average autosomal  $F_{\text{ST}} = 0.223$  (Fig. 1b and Supplementary Table 5). All  $F_{\text{ST}}$  regions are characterized by low levels of heterozygosity in either dog or wolf (although all do not pass the  $Z(H_p) < -5$  threshold), indicating that the two statistics detect the same events (Methods and Supplementary Discussion, sections 2 and 3). In total, 36 unique autosomal candidate domestication regions (CDRs) containing 122 genes were identified by the two approaches combined (Supplementary Table 6 and Fig. 1b, c). None of these regions overlaps those of a previous genotype-based study<sup>5</sup> (Supplementary Discussion, section 3), stressing the importance of identifying domestication regions directly by sequencing or by comprehensively ascertaining SNPs in wild ancestors before genotyping.

We searched for significantly overrepresented gene ontology terms among genes in autosomal CDRs and identified 25 categories, representing several groups of interrelated terms (Table 1 and Supplementary Table 7), none of which was indicated in a separate analysis of selection in wolf (Supplementary Discussion, section 8). The most conspicuous cluster (11 terms) relates to the term ‘nervous system development’. The eight genes belonging to this category (Supplementary Tables 7 and 8) include *MBP*, *VWC2*, *SMO*, *TLX3*, *CYFIP1* and *SH3GL2*, of which several affect developmental signalling and synaptic strength and plasticity<sup>12–16</sup>. We surveyed published literature and identified 11 additional CDR genes with central nervous system function (Supplementary Table 9), adding to a total of 19 CDRs that contain brain genes. These findings support the hypothesis that selection for altered behaviour was important during dog domestication and that mutations affecting developmental genes may underlie these changes<sup>7</sup>.

The gene ontology analysis also pinpoints two genes involved in the binding of sperm and egg: *ZPBP* encodes the zona pellucida binding protein that mediates binding of sperm to the zona pellucida glycoprotein layer (ZP) of the egg, and *ZP2* codes for one of the proteins that make up ZP itself. In addition, a CDR on chromosome 6 encompasses *PDILT* that also affects binding of sperm to ZP<sup>17</sup>, altogether indicating that sperm competition may have been an important evolutionary force during dog domestication<sup>18</sup>.

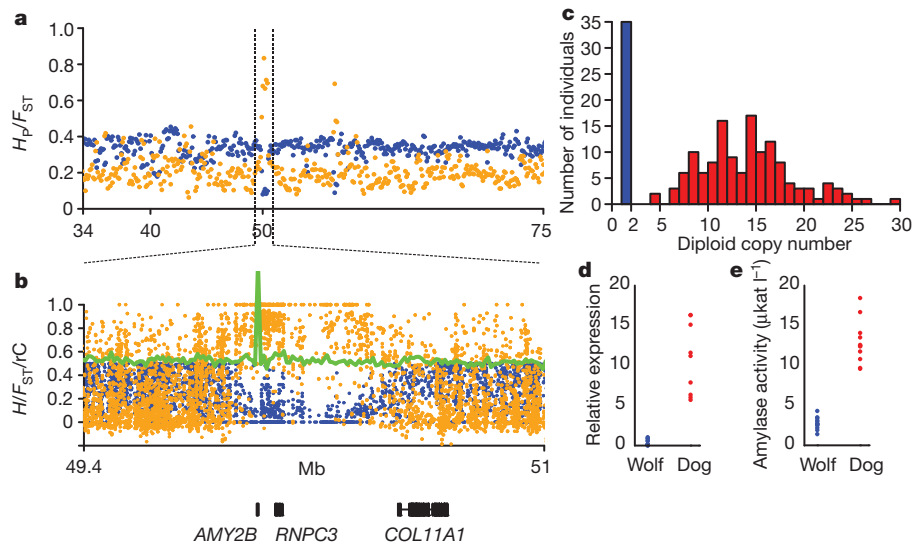
Overrepresented terms ‘starch metabolic process’, ‘digestion’ and ‘fatty acid metabolism’ include genes involved in starch digestion (*MGAM*) and glucose uptake (*SGLT1*), as well as a candidate gene for

**Table 1 | Enriched gene ontology terms among CDR genes**

Gene ontology term	$P_{\text{FDR}}$ value	Gene count
Regulation of neuron differentiation	0.005	3 (26)
Multicellular organismal process	0.005	21 (3,822)
Digestion	0.008	4 (95)
Neuron differentiation	0.010	5 (210)
Regulation of molecular function	0.011	8 (671)
Central nervous system development	0.013	5 (235)
Regulation of developmental process	0.013	5 (236)
Generation of neurons	0.013	5 (242)
Nervous system development	0.013	8 (716)
Binding of sperm to zona pellucida	0.015	2 (12)
Sperm-egg recognition	0.015	2 (12)
Neurogenesis	0.015	5 (262)
Cell-cell recognition	0.019	2 (14)
Regulation of catalytic activity	0.020	7 (605)
Regulation of hydrolase activity	0.026	5 (307)
Fatty acid metabolic process	0.031	4 (191)
System development	0.034	11 (1,605)
Regulation of GTPase activity	0.039	4 (211)
Anatomical structure development	0.039	12 (2,005)
Intramembranous ossification	0.039	1 (1)
Quinolinate metabolic process	0.039	1 (1)
Starch metabolic process	0.039	1 (1)
Starch catabolic process	0.039	1 (1)
Glucocorticoid catabolic process	0.039	1 (1)
Cell development	0.039	9 (1,242)

Enriched terms are colour-coded to reflect relatedness in the ontology or functional proximity. Blue, nervous system development; green, sperm-egg recognition; grey, regulation of molecular function; orange, digestion. For each term, gene count shows number of genes in CDRs relative to total number of annotated genes (in parentheses).





**Figure 2 | Selection for increased amylase activity.** **a**, Pooled heterozygosity,  $H_P$  (blue), and average fixation index,  $F_{ST}$  (orange), plotted for 200-kb windows across a chromosome 6 region harbouring *AMY2B*. **b**, Heterozygosity,  $H$  (blue), and fixation index,  $F_{ST}$  (orange), for single SNPs in the selected region. Dog relative to wolf coverage,  $rC$  (green line), indicates increase in *AMY2B* copy

insulin resistance (*ACSM2A*) that initiates the fatty acid metabolism<sup>19</sup>. A total of 6 CDRs harbour 10 genes with functions related to starch and fat metabolism (Supplementary Table 10). We propose that genetic variants within these genes may have been selected to aid adaptation from a mainly carnivorous diet to a more starch rich diet during dog domestication.

The breakdown of starch in dogs proceeds in three stages: (1) starch is first cleaved to maltose and other oligosaccharides by alpha-amylase in the intestine; (2) the oligosaccharides are subsequently hydrolysed by maltase-glucoamylase<sup>20</sup>, sucrase and isomaltase to form glucose; and (3) finally, glucose is transported across the plasma membrane by brush border protein SGLT1<sup>21</sup>. Here we present evidence for selection on all three stages of starch digestion during dog domestication.

Whereas humans have acquired amylase activity in the saliva<sup>22</sup> via an ancient duplication of the pancreatic amylase gene, dogs only express amylase in the pancreas<sup>23</sup>. In dogs the *AMY2B* gene, encoding the alpha-2B-amylase, resides in a 600-kb CDR on chromosome 6 with  $Z(H_P)$  and  $Z(F_{ST})$  scores of  $-4.60$  and  $7.16$ , respectively (Figs 1 and 2a). Interestingly, an 8-kb sequence spanning the *AMY2B* locus showed a several-fold increase in aligned read depth in dog relative to wolf (Fig. 2b), suggestive of a copy number change. Formal comparisons of regional and local pool coverage, and wolf and dog coverage (Methods), respectively, also suggest a substantial increase in copy numbers in all dog pools compared to wolf at this locus (Supplementary Discussion, section 5).

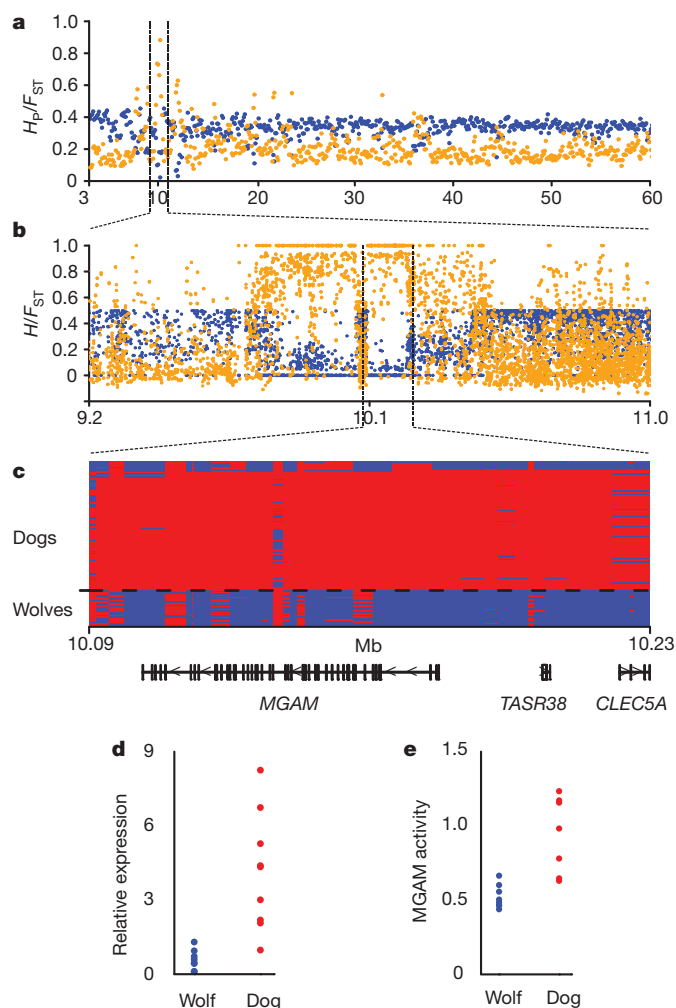
We confirmed this CNV by quantifying *AMY2B* copy numbers in 136 dogs and 35 wolves (Supplementary Table 11) using real-time quantitative PCR (qPCR). Whereas all wolves tested carried only 2 copies ( $2N = 2$ ), diploid copy numbers in dog ranged from 4 to 30 ( $P < 0.001$ , Wilcoxon) (Fig. 2c), corresponding to a remarkable 7.4-fold average increase in dog *AMY2B* copy numbers. To assess whether this change correspond to a difference in amylase activity, we first compared *AMY2B* gene expression in pancreas from dog ( $n = 9$ ) and wolf ( $n = 12$ ) and noted a 28-fold higher average expression in dog ( $P < 0.001$ , Wilcoxon, Fig. 2d). We then quantified amylase activity in frozen serum (Fig. 2e) and found a 4.7-fold higher activity in dog ( $9.6\text{--}18.4 \mu\text{kat l}^{-1}$  ( $n = 12$ )) relative to wolf ( $1.4\text{--}4.3 \mu\text{kat l}^{-1}$  ( $n = 13$ )) ( $P < 0.001$ , Wilcoxon). Similar results were obtained in comparisons of a limited number of fresh samples (Supplementary Tables 12 and 13). The change in *AMY2B* gene copy

number in dog. Genes in the region are shown below panel **b**. **c**, Histogram showing the distribution of diploid amylase copy number in wolf ( $n = 35$ ) (blue) and dog ( $n = 136$ ) (red). **d**, Amylase messenger RNA expression levels in pancreas of wolf ( $n = 12$ ) and dog ( $n = 9$ ). **e**, Amylase activity in serum from wolf ( $n = 13$ ) and dog ( $n = 12$ ).

number together with a correlated increase in both expression level ( $\rho = 0.84$ ,  $P < 0.0001$ , Spearman) (Supplementary Fig. 3) and enzyme activity ( $\rho = 0.63$ ,  $P < 0.01$ , Spearman) (Supplementary Fig. 4) indicates that duplications of the alpha-amylase locus conferred a selective advantage to early dogs by causing an increase in amylase activity.

Maltase-glucoamylase is responsible for the second step in the breakdown of starch, catalysing the hydrolysis of maltose to glucose<sup>20</sup>. No copy number changes were observed in the *MGAM* locus so we decided to study haplotype diversity across the region to facilitate the identification of causal variants. We genotyped 47 randomly selected SNPs in 71 dogs representing 38 diverse breeds and 19 wolves of worldwide distribution (referred to as 'the reference panel', Supplementary Table 14). Sixty-eight of the seventy-one dogs tested carried at least one copy of a 124-kb long haplotype spanning the entire *MGAM* and a small neighbouring locus encoding the bitter taste mediating taste receptor 2 member 38 (*TASR38*) (Fig. 3a–c). Whereas none of the wolves carried the selected haplotype, 55 dogs were homozygous for it, 13 were heterozygous and only three dogs lacked it (2 West Highland White Terriers and 1 Chinese Crested Dog). This high degree of haplotype differentiation between dog and wolf (average  $F_{ST}$  for genotyped SNPs = 0.75) indicates that this haplotype may harbour genetic variation of selective advantage to dogs (Supplementary Discussion, sections 3 and 6).

We identified several candidate mutations within *MGAM* that may have been targeted by selection in this region (Supplementary Table 15). First a conservative amino acid substitution located in the duplicated trefoil domain of *MGAM* (residue 1001) is nearly fixed for isoleucine in wolf and for valine in dogs. Eleven out of fourteen mammals have valine at this position, whereas the omnivorous rat, and the insectivorous hedgehog and short-tailed opossum, carry isoleucine like the wolf (Supplementary Table 16). Second, another conservative substitution, methionine to valine, located in the beta sheet of the maltase enzyme (residue 797), is segregating in wolf but fixed for methionine in dog. The insectivorous hedgehog and common shrew are the only mammals without methionine at this evolutionarily conserved position (Supplementary Table 17) and *in silico* modelling using the SDM-server indicates that a change from methionine to valine at this residue is destabilizing<sup>24</sup>. Third, a fixed two-base-pair deletion in dog disrupts the stop codon, thereby extending the carboxy-terminal end of dog



**Figure 3 | Selection is associated with increased maltase activity.** **a**, Pooled heterozygosity,  $H_P$  (blue), and average fixation index,  $F_{ST}$  (orange), plotted for 200-kb windows across a chromosome 16 region harbouring *MGAM*. **b**, Heterozygosity,  $H$  (blue), and fixation index,  $F_{ST}$  (orange), for single SNPs in the selected region. **c**, Haplotypes inferred from genotyping of 47 SNPs across the *MGAM* locus in 71 dogs and 19 wolves (red and blue colour are major and minor dog allele, respectively). Genes in the genotyped region are shown below panel **c**. **d**, *MGAM* mRNA expression levels in pancreas of wolf ( $n = 8$ ) and dog ( $n = 9$ ). **e**, *MGAM* activity in serum from wolf ( $n = 8$ ) and dog ( $n = 7$ ).

*MGAM* by two amino acids: asparagine and phenylalanine. In 32 mammals studied only herbivores (rabbit, pika, alpaca and cow) and omnivores (mouse lemur and rat) share an extension like that seen in dog (Supplementary Table 18). A fourth candidate mutation in intron 37 affects a predicted binding site for the glucose metabolism regulator NR4A2 protein<sup>25</sup> by shifting the wolf sequence away from the canonical NR4A2-binding motif. Three out of four mammals with the wolf allele at this site rely heavily on insects or fish for their nutritional requirements (Supplementary Table 19).

To decipher whether the candidate mutations act primarily on expression or protein activity we examined *MGAM* expression in pancreas and the resulting enzymatic activity in serum. Dogs showed a ~12-fold higher expression ( $P < 0.001$ , Wilcoxon,  $n_{\text{DOG}} = 9$ ,  $n_{\text{WOLF}} = 8$ ) (Fig. 3d) and a ~twofold increase in maltose to glucose turnover compared to wolves (average glucose produced in dogs:  $0.94 \Delta A_{570 \text{ nm}}$  (0.64–1.23,  $n = 7$ ) and wolves:  $0.52 \Delta A_{570 \text{ nm}}$  (0.44–0.66,  $n = 8$ ),  $P = 0.0012$ , Wilcoxon) (Fig. 3e). Although we cannot rule out that diet-induced plasticity contributed to this difference<sup>26</sup>, our results indicate that the mutation affecting a NR4A2-binding site or another unknown variant probably affect the expression of *MGAM*. Selection may thus clearly have

led to increased *MGAM* expression, but we cannot rule out that the strong selection affecting this locus may have favoured the accumulation of protein-coding changes on the same haplotype. Similar scenarios have been seen for white coat colour in dogs and pigs, where repeated selection for additional mutations has resulted in an allelic series of white spotting at the *MITF* and *KIT* loci, respectively<sup>27</sup>.

Once starch has been digested to glucose it is absorbed through the luminal plasma membrane of the small intestine by the sodium/glucose cotransporter 1 (SGLT1)<sup>21</sup>. To benefit from an increased capacity to digest starch, dogs would therefore be expected to show a parallel increase in glucose uptake. A CDR on chromosome 26 (Supplementary Fig. 5a, b) encompasses *SGLT1* and a gene (*SGLT3*) encoding the glucose-sensing sodium/glucose cotransporter 3 protein<sup>28</sup>. To characterize the haplotype diversity we genotyped 48 randomly chosen SNPs across this CDR in the reference panel and identified a 50.5-kb region, spanning the 3' section of *SGLT1* as well as the 3' end of *SGLT3*, that is highly divergent between dog and wolf (Supplementary Fig. 5c). In this region all dogs tested were carriers of a particular haplotype, for which 63 were homozygous and eight heterozygous. This contrasts to 19 wolves where a single individual carried one copy of the haplotype. Based on the high haplotype differentiation (average  $F_{ST}$  for 18 SNPs in 50.5-kb haplotype = 0.81) it is likely that *SGLT1* and its 3' region represents an additional dog domestication locus.

The 50.5-kb region includes a conservative isoleucine to valine substitution in *SGLT1* (residue 244) that affects a loop facing the extracellular side of the luminal membrane (Supplementary Table 15). Heterologous expression analysis<sup>29</sup> shows that glycosylation at a nearby site (residue 248) affects glucose transport, indicating that it is possible that dogs acquired improved glucose uptake as a result of the observed substitution. In addition, we see only non-significant differences in *SGLT1* expression in pancreas of dog ( $n = 9$ ) and wolf ( $n = 4$ ) ( $P = 0.39$ , Wilcoxon) (Supplementary Fig. 6), indicating that selection primarily targeted a structural rather than regulatory mutation in *SGLT1*.

In conclusion, we have presented evidence that dog domestication was accompanied by selection at three genes with key roles in starch digestion: *AMY2B*, *MGAM* and *SGLT1*. Our results show that adaptations that allowed the early ancestors of modern dogs to thrive on a diet rich in starch, relative to the carnivorous diet of wolves, constituted a crucial step in early dog domestication. This may suggest that a change of ecological niche could have been the driving force behind the domestication process, and that scavenging in waste dumps near the increasingly common human settlements during the dawn of the agricultural revolution may have constituted this new niche<sup>6</sup>. In light of previous results describing the timing and location of dog domestication, our findings may suggest that the development of agriculture catalysed the domestication of dogs.

The results presented here demonstrate a striking case of parallel evolution whereby the benefits of coping with an increasingly starch-rich diet during the agricultural revolution caused similar adaptive responses in dog and human<sup>30</sup>. This emphasizes how insights from dog domestication may benefit our understanding of human recent evolution and disease. Finally, by understanding the genetic basis of adaptive traits in dogs we have come closer to unlocking the potential in dog and wolf comparisons to decipher the genetics of behaviour.

## METHODS SUMMARY

**Sequencing.** We pooled genomic DNA from 12 individuals before mate-pair library construction and sequencing on the AB SOLiD system, version 3, according to standard manufacturer protocols. Sequencing reads were aligned to the CanFam 2.0 reference sequence using the Bioscope 1.1 software.

**Selection analyses.** We identified variable sites in data combined from all pools and required a minimum of three reads supporting an alternative allele to call a SNP. We used allele counts at variable sites to identify signals of selection in 200-kb windows using two approaches: for each window we calculated (1) the average pooled heterozygosity,  $H_P$  (ref. 10), and (2) the average fixation index,  $F_{ST}$ , between dog and wolf. Putatively selected regions were located by extracting



windows from the extreme tails of the  $Z$ -transformed  $H_P$  and  $F_{ST}$  distributions by applying a threshold of 5 standard deviations.

**Functional assays.** We used multiplex TaqMan assays and SYBR Green real-time PCR to quantify CNVs and gene expression, respectively. Serum amylase activity was analysed using an Architect e400 instrument and serum maltase activity was quantified based on the amount of maltose to glucose turnover.

**Full Methods** and any associated references are available in the online version of the paper.

Received 1 July; accepted 11 December 2012.

Published online 23 January 2013.

- Ovodov, N. D. *et al.* A 33,000-year-old incipient dog from the Altai mountains of Siberia: evidence of the earliest domestication disrupted by the last glacial maximum. *PLoS ONE* **6**, e22821 (2011).
- Davis, S. J. M. & Valla, F. R. Evidence for domestication of the dog 12,000 years ago in the Natufian of Israel. *Nature* **276**, 608–610 (1978).
- Skoglund, P., Götherström, A. & Jakobsson, M. Estimation of population divergence times from non-overlapping genomic sequences: examples from dogs and wolves. *Mol. Biol. Evol.* **28**, 1505–1517 (2011).
- Pang, J. F. *et al.* mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Mol. Biol. Evol.* **26**, 2849–2864 (2009).
- vonHoldt, B. M. *et al.* Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **464**, 898–902 (2010).
- Coppinger, R. & Coppinger, L. *Dogs: a Startling New Understanding of Canine Origin, Behaviour and Evolution* (Scribner, 2001).
- Hare, B., Wobber, V. & Wrangham, R. The self-domestication hypothesis: evolution of bonobo psychology is due to selection against aggression. *Anim. Behav.* **83**, 573–585 (2012).
- Belyaev, D. K. Destabilizing selection as a factor in domestication. *J. Hered.* **70**, 301–308 (1979).
- Fang, M., Larson, G., Ribeiro, H. S., Li, N. & Andersson, L. Contrasting mode of evolution at a coat color locus in wild and domestic pigs. *PLoS Genet.* **5**, e1000341 (2009).
- Rubin, C. J. *et al.* Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* **464**, 587–591 (2010).
- Lindblad-Toh, K. *et al.* Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819 (2005).
- Koike, N. *et al.* Brorin, a novel secreted bone morphogenetic protein antagonist, promotes neurogenesis in mouse neural precursor cells. *J. Biol. Chem.* **282**, 15843–15850 (2007).
- Cheng, L. *et al.* *Tlx3* and *Tlx1* are post-mitotic selector genes determining glutamatergic over GABAergic cell fates. *Nature Neurosci.* **7**, 510–517 (2004).
- Napoli, I. *et al.* The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. *Cell* **134**, 1042–1054 (2008).
- Weston, M. C., Nehring, R. B., Wojcik, S. M. & Rosenmund, C. Interplay between VGLUT isoforms and endophilin A1 regulates neurotransmitter release and short-term plasticity. *Neuron* **69**, 1147–1159 (2011).
- Varga, Z. M. *et al.* Zebrafish smoothened functions in ventral neural tube specification and axon tract formation. *Development* **128**, 3497–3509 (2001).
- Tokuhiro, K., Ikawa, M., Benham, A. M. & Okabe, M. Protein disulfide isomerase homolog PDILT is required for quality control of sperm membrane protein ADAM3 and male fertility. *Proc. Natl Acad. Sci. USA* **109**, 3850–3855 (2012).
- Gardner, A. J. & Evans, J. P. Mammalian membrane block to polyspermy: new insights into how mammalian eggs prevent fertilisation by multiple sperm. *Reprod. Fertil. Dev.* **18**, 53–61 (2006).
- Boomgaarden, I., Vock, C., Klapper, M. & Doring, F. Comparative analyses of disease risk genes belonging to the acyl-CoA synthetase medium-chain (ACSM) family in human liver and cell lines. *Biochem. Genet.* **47**, 739–748 (2009).
- Nichols, B. L. *et al.* The maltase-glucoamylase gene: common ancestry to sucrase-isomaltase with complementary starch digestion activities. *Proc. Natl Acad. Sci. USA* **100**, 1432–1437 (2003).
- Wright, E. M., Loo, D. D. F. & Hirayama, B. A. Biology of human sodium glucose transporters. *Physiol. Rev.* **91**, 733–794 (2011).
- Meisler, M. H. & Ting, C. N. The remarkable evolutionary history of the human amylase genes. *Crit. Rev. Oral Biol. Med.* **4**, 503–509 (1993).
- Simpson, J. W., Doxey, D. L. & Brown, R. Serum isoamylase values in normal dogs and dogs with exocrine pancreatic insufficiency. *Vet. Res. Commun.* **8**, 303–308 (1984).
- Worth, C. L., Preissner, R. & Blundell, T. L. SDM—a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res.* **39**, W215–W222 (2011).
- Pei, L. *et al.* NR4A orphan nuclear receptors are transcriptional regulators of hepatic glucose metabolism. *Nature Med.* **12**, 1048–1055 (2006).
- Mochizuki, K., Honma, K., Shimada, M. & Goda, T. The regulation of jejunal induction of the maltase-glucoamylase gene by a high-starch/low-fat diet in mice. *Mol. Nutr. Food Res.* **54**, 1445–1451 (2010).
- Andersson, L. Studying phenotypic evolution in domestic animals: a walk in the footsteps of Charles Darwin. *Cold Spring Harb. Symp. Quant. Biol.* **74**, 319–325 (2009).
- Diez-Sampedro, A. *et al.* A glucose sensor hiding in a family of transporters. *Proc. Natl Acad. Sci. USA* **100**, 11753–11758 (2003).
- Hediger, M. A., Mendlein, J., Lee, H. S. & Wright, E. M. Biosynthesis of the cloned intestinal Na<sup>+</sup> glucose cotransporter. *Biochim. Biophys. Acta* **1064**, 360–364 (1991).
- Perry, G. H. *et al.* Diet and the evolution of human amylase gene copy number variation. *Nature Genet.* **39**, 1256–1260 (2007).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank Järvzoo, Nordens ark and the Canine Biobank at Uppsala University and the Swedish University of Agricultural Sciences for providing samples, Uppsala Genomics Platform at SciLifeLab Uppsala for generating the resequencing data, the UPPNEX platform for assisting with computational infrastructure for data analysis and the Broad Institute Genomics Platform for validation genotyping. The project was funded by the SSF, the Swedish Research Council, the Swedish Research Council Formas, Uppsala University and a EURL to K.L.-T. funded by the ESF supporting also E.A.; K.M. was funded by the Higher Education Commission, Pakistan.

**Author Contributions** K.L.-T. and Å.H. designed the study. K.L.-T. and E.A. oversaw the study. M.-L.A. coordinated and performed the majority of the sample collecting and O.L. and J.M.A. provided samples of critical importance. E.A. performed the SNP detection and selection analyses; A.R. identified candidate causative mutations and analysed haplotypes in CDRs; K.M. detected CNVs bioinformatically; M.T.W. performed phylogenetic analysis and analysed the Canine HD-array data; A.R. performed the maltase activity assay; M.-L.A. validated CNVs and quantified mRNA expression of candidate genes; M.P. performed validation SNP genotyping; E.A., A.R., M.-L.A. and K.L.-T. interpreted the data; E.A. and K.L.-T. wrote the paper with input from the other authors.

**Author Information** Sequence reads are available under the accession number SRA061854 (NCBI Sequence Read Archive). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.A. (Erik.Axelsson@imbim.uu.se) and K.L.-T. (kersli@broadinstitute.org).

## METHODS

**DNA extraction.** DNA was extracted from tissue using Qiagen tissue DNA extraction kits or from EDTA blood using either manual salt precipitation or the QIASymphony DNA Midi kit (Qiagen) on the QIASymphony robot (Qiagen).

**Sequencing.** We pooled DNA from 12 individuals per pool before mate-pair library construction and sequencing on the AB SOLiD system, version 3, according to standard manufacturer protocols (Applied Biosystems). Sequencing reads were aligned to the CanFam 2.0 reference sequence using the Bioscope 1.1 software. We removed duplicated (<http://picard.sourceforge.net>) and poorly mapped reads (mapping quality <20 in Samtools)<sup>31</sup> and retained only uniquely mapped reads for further analyses.

**SNP detection.** We searched for variable sites in data combined from all pools (including wolf) to increase sensitivity to rare alleles. We required a minimum of three reads supporting an alternative allele to call a SNP, and applied a further filtering step implemented in samtools.pl varFilter (settings: -Q25 -q10 -d3 -D120 -G25 -w10 -N2 -l30) to ensure a high call accuracy that is largely unaffected by, for example, paralogous sequence variants. We called genotypes for all SNPs in all dog pools and the single wolf pool by counting sequencing reads supporting the reference and variant allele, given a minimum base quality of 20, to estimate allele frequencies in the dog and wolf populations. A random selection representing 25% of the sequencing reads from pools 4 and 5 were included in this process to achieve unbiased allele frequency estimates.

**Selection analyses.** Allele counts and allele frequencies at all identified variable sites were used to search the dog genome for regions that may have been affected by selection during the early phase of dog domestication using two complementary approaches. First we calculated the average pooled heterozygosity ( $H_p$ ) in 200-kb windows sliding 100 kb at a time, for all five dog pools combined, and in the single wolf pool separately, following the methodology described in ref. 10. Briefly, this method sums all minor and major allele counts, respectively, at all variable sites within a window, and estimates the heterozygosity based on the combined allele counts for the entire window. The advantage of this method over calculating a simple arithmetic mean of all single-site heterozygosity estimates is that it accounts for variable sequence coverage across the window. To avoid spurious selection signals we discarded 49 out of 21,927 windows containing fewer than 10 informative sites from both this and the subsequent  $F_{ST}$  analysis. We Z-transformed the resultant distribution of  $H_p$  scores and extracted putatively selected windows in the extreme tail of the distribution by applying a  $Z(H_p) < -5$  cut-off.

Second we calculated  $F_{ST}$  values between dog and wolf for individual SNPs using a method that adjusts for sample size differences<sup>32</sup>. We averaged  $F_{ST}$  values across 200-kb windows, sliding 100 kb at a time and Z-transformed the resultant distribution. Putative selection targets were extracted from the extreme tail of the distribution by applying a  $Z(F_{ST}) > 5$  cut-off, and attributed to selection in dog if the corresponding  $Z(H_p)_{DOG} < Z(H_p)_{WOLF}$ , to selection in wolf if  $Z(H_p)_{WOLF} < Z(H_p)_{DOG}$  (three regions), and to selection in both taxa if  $Z(H_p)_{WOLF} < -4$  and  $Z(H_p)_{WOLF} < -4$ .

**Gene ontology analysis.** We used the Ensembl gene annotations to identify genes residing within regions extending 100 kb up- and downstream of CDRs to include potential effects of regulatory changes on loci at some distance, and to reduce the risk of excluding the outermost portions of the selected haplotypes by using sliding windows of fixed size. We tested for enrichment of gene ontology terms (GOa-human) assigned to the subset of these CDR genes for which human orthology could be established (79 out of 122) using the GStat program<sup>33</sup>.

**Genotyping validation.** We designed an iPLEX assay targeting 124 SNPs located in CDRs showing a high degree of homozygosity or population differentiation. A total of 71 dogs, representing 38 different breeds, and 19 wolves (Supplementary Table 14) were genotyped using standard protocols provided by the manufacturer (Sequenome). Haplotypes were phased using fastPHASE<sup>34</sup>.

**qPCR CNV detection.** We quantified DNA copy number variation using Multiplex TaqMan assays containing primers and probes (Supplementary Table 20)

matching both the target and reference sequence (housekeeping gene *C7orf28b*) according to the manufacturer's protocol. All reactions were run in triplicate and data was analysed using the CopyCaller software (Applied Biosystems). Copy numbers for each target were normalized to the same wolf to account for inter-plate variability.

**qPCR expression analyses.** Pancreatic tissue samples from dogs and wolves where collected post mortem, stored in RNAlater at 4 °C for 24 h and subsequently freeze-stored at -80 °C. We used TRIzol to isolate RNA from these samples, followed by complementary DNA synthesis using the Advantage RT for PCR kit according to the manufacturers' protocols (Life Technologies and Clontech, respectively). We designed exonic primers (Supplementary Table 21) and quantified the amount of cDNA using SYBR Green real-time PCR (Applied Biosystems) on a 7900HT Fast real time PCR system (Applied Biosystems) and analysed the data using the qbasePLUS (Biogazelle) software according to the  $\Delta\Delta C_T$  method. All reactions where run in triplicate and normalized by comparisons to housekeeping genes *RPL32* and *RPL13A*.

**Amylase activity.** Peripheral EDTA and serum blood samples where collected from dogs and both captive and free-ranging wolves. Serum amylase activity was analysed at the Clinical Pathology service (Swedish Agricultural University) using an Architect e400 instrument (Abbott Laboratories), except for 8 serum samples (Supplementary Table 13) which were run on a VetScan instrument (Abaxis).

**Maltase activity.** Maltase activity was assayed according to the principle outlined in ref. 35, whereby a known amount of maltose substrate is added to serum and the resultant glucose produced is measured as the change in absorbance after five minutes ( $\Delta A_{570\text{nm}}$ ). We used reagents from the ab83388 Maltose assay kit (Abcam) and serum sampled as described above. For each individual, glucose residuals were measured in duplicate and maltase assays were performed in triplicate.

**Indel calling.** We used Bioscope 1.1 to call small insertions and deletions in each pool separately. We then combined the results of all pools and extracted a set of high confident indels by requiring that indels were supported by at least three sequencing reads.

**CNV detection.** Four methods were used to detect structural variation in the dog genome. We searched for deviations in insert size using the large indel tool implemented in Bioscope1.1. We compared the coverage depth between the pooled samples using CNVseq<sup>36</sup> and the Fixed deletions method<sup>10</sup> and finally identified regions in which the coverage depth deviated from the pool average using CNVnator<sup>37</sup>. Methods relying on comparisons of sequence coverage between pools always used the wolf as reference pool.

**Ethics.** All animals contributing tissue samples to this study died for other reasons than participating in this study. All dog samples were taken with the owners consent. The sampling conformed to the decision of the Swedish Animal Ethical Committee (no. C62/10) and the Swedish Animal Welfare Agency (no.31-1711/10).

31. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
32. Weir, B. S. & Cockerham, C. C. Estimating  $F$ -statistics for the analysis of population-structure. *Evolution* **38**, 1358–1370 (1984).
33. Beissbarth, T. & Speed, T. P. GStat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* **20**, 1464–1465 (2004).
34. Scheet, P. & Stephens, M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* **78**, 629–644 (2006).
35. Dahlqvist, A. Method for assay of intestinal disaccharidases. *Anal. Biochem.* **7**, 18–25 (1964).
36. Xie, C. & Tammi, M. T. CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinformatics* **10**, 80 (2009).
37. Abyzov, A., Urban, A. E., Snyder, M. & Gerstein, M. CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**, 974–984 (2011).



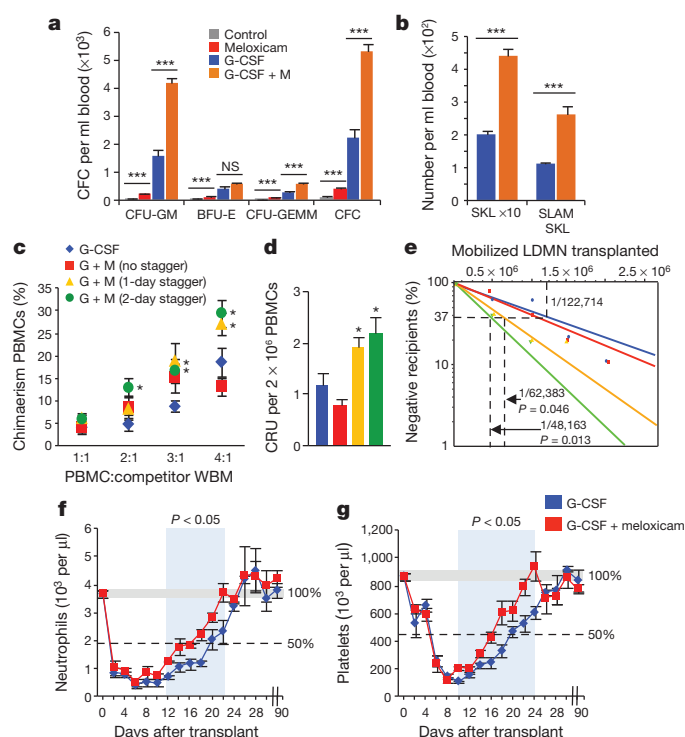
# Differential stem- and progenitor-cell trafficking by prostaglandin E<sub>2</sub>

Jonathan Hoggatt<sup>1,2</sup>, Khalid S. Mohammad<sup>3\*</sup>, Pratibha Singh<sup>1\*</sup>, Amber F. Hoggatt<sup>1,4</sup>, Brahmananda R. Chitteti<sup>5</sup>, Jennifer M. Speth<sup>1</sup>, Peirong Hu<sup>1</sup>, Bradley A. Poteat<sup>5</sup>, Kayla N. Stilger<sup>1</sup>, Francesca Ferraro<sup>2</sup>, Lev Silberstein<sup>2</sup>, Frankie K. Wong<sup>2</sup>, Sherif S. Farag<sup>5</sup>, Magdalena Czader<sup>6</sup>, Ginger L. Milne<sup>7</sup>, Richard M. Breyer<sup>8</sup>, Carlos H. Serezani<sup>1</sup>, David T. Scadden<sup>2</sup>, Theresa A. Guise<sup>3</sup>, Edward F. Srouf<sup>1,5</sup> & Louis M. Pelus<sup>1</sup>

To maintain lifelong production of blood cells, haematopoietic stem cells (HSCs) are tightly regulated by inherent programs and extrinsic regulatory signals received from their microenvironmental niche. Long-term repopulating HSCs reside in several, perhaps overlapping, niches that produce regulatory molecules and signals necessary for homeostasis and for increased output after stress or injury<sup>1–5</sup>. Despite considerable advances in the specific cellular or molecular mechanisms governing HSC–niche interactions, little is known about the regulatory function in the intact mammalian haematopoietic niche. Recently, we and others described a positive regulatory role for prostaglandin E<sub>2</sub> (PGE<sub>2</sub>) on HSC function *ex vivo*<sup>6,7</sup>. Here we show that inhibition of endogenous PGE<sub>2</sub> by non-steroidal anti-inflammatory drug (NSAID) treatment in mice results in modest HSC egress from the bone marrow. Surprisingly, this was independent of the SDF-1–CXCR4 axis implicated in stem-cell migration. Stem and progenitor cells were found to have differing mechanisms of egress, with HSC transit to the periphery dependent on niche attenuation and reduction in the retentive molecule osteopontin. Haematopoietic grafts mobilized by NSAIDs had superior repopulating ability and long-term engraftment. Treatment of non-human primates and healthy human volunteers confirmed NSAID-mediated egress in other species. PGE<sub>2</sub> receptor knockout mice demonstrated that progenitor expansion and stem/progenitor egress resulted from reduced E-prostanoid 4 (EP4) receptor signalling. These results not only uncover unique regulatory roles for EP4 signalling in HSC retention in the niche, but also define a rapidly translatable strategy to enhance transplantation therapeutically.

Mice were treated with the prototypical NSAID indomethacin (Supplementary Fig. 1a) to reduce endogenous PGE<sub>2</sub> production, resulting in a significant increase in haematopoietic progenitor cells (HPCs) in the peripheral blood that was not accompanied by an increase in the number of white blood cells (Supplementary Fig. 1b, c). This HPC increase but unaltered white blood cell count probably accounts for the lack of previous detection of this observation, despite decades of clinical NSAID use. No increase in HPC egress was seen in mice treated with the lipoxigenase inhibitor baicalin, suggesting a cyclooxygenase (COX) pathway-specific effect. Co-administration of indomethacin with the clinically used mobilizing agent granulocyte colony-stimulating factor (G-CSF) significantly enhanced (~twofold) HPC mobilization (Supplementary Fig. 1b). NSAIDs with varying degrees of COX-1 and COX-2 selectivity demonstrated significant mobilization with indomethacin, aspirin, ibuprofen and meloxicam (Supplementary Fig. 2). Meloxicam inhibits both COX-1 and COX-2 in the bone marrow microenvironment (Supplementary Fig. 3), and when compared to other dual

inhibitors it has a reduced incidence of gastrointestinal discomfort<sup>8</sup> and inhibition of platelet aggregation<sup>9</sup>. Therefore, meloxicam was used in most of the studies. We did not extensively test the differential roles of

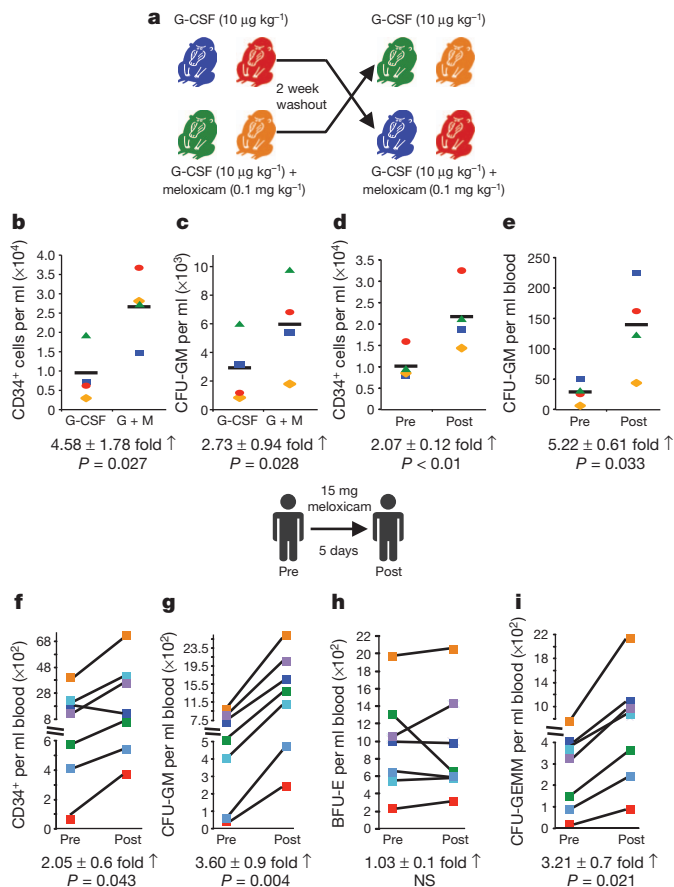


**Figure 1 | NSAIDs mobilize haematopoietic stem and progenitor cells.**

**a, b**, Meloxicam enhances mobilization of HPCs (**a**) and HSCs (**b**) into blood ( $n = 4–5$  mice per group per experiment; three experiments). BFU-E, erythroid progenitors (burst-forming unit–erythrocyte); CFC, colony-forming cells; CFU-GEMM, multipotential progenitor cells (colony-forming unit–granulocyte, erythrocyte, macrophage, megakaryocyte); M, meloxicam. **c**, Chimaerism. PBMCs, peripheral blood mononuclear cells; WBM, whole bone marrow cells. **d, e**, Competitive repopulating units (CRU) (**d**), and LT-HSC frequency (Poisson distribution) (**e**), 36 weeks after limiting dilution competitive transplants of PBMCs from mice treated with G-CSF and combination regimens ( $n = 8$  mice per group, assayed individually). Mice were treated with G-CSF or a staggered regimen of G-CSF plus meloxicam, and PBMCs were transplanted into lethally irradiated mice. **f, g**, Neutrophil (**f**) and platelet (**g**) recovery were monitored for 90 days. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; unpaired two-tailed *t*-test. All data are mean  $\pm$  s.e.m.

<sup>1</sup>Microbiology and Immunology, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA. <sup>2</sup>Stem Cell and Regenerative Biology, Harvard University/Harvard Stem Cell Institute/Harvard Medical School/Center for Regenerative Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. <sup>3</sup>Medicine/Endocrinology, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA. <sup>4</sup>Biologic Resources Laboratory, University of Illinois at Chicago, Chicago, Illinois 60612, USA. <sup>5</sup>Medicine/Division of Hematology and Oncology, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA. <sup>6</sup>Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA. <sup>7</sup>Eicosanoid Core Laboratory, Division of Clinical Pharmacology, Vanderbilt University, Nashville, Tennessee 37232, USA. <sup>8</sup>Division of Nephrology and Hypertension, Vanderbilt University, Nashville, Tennessee 37232, USA.

\*These authors contributed equally to this work.

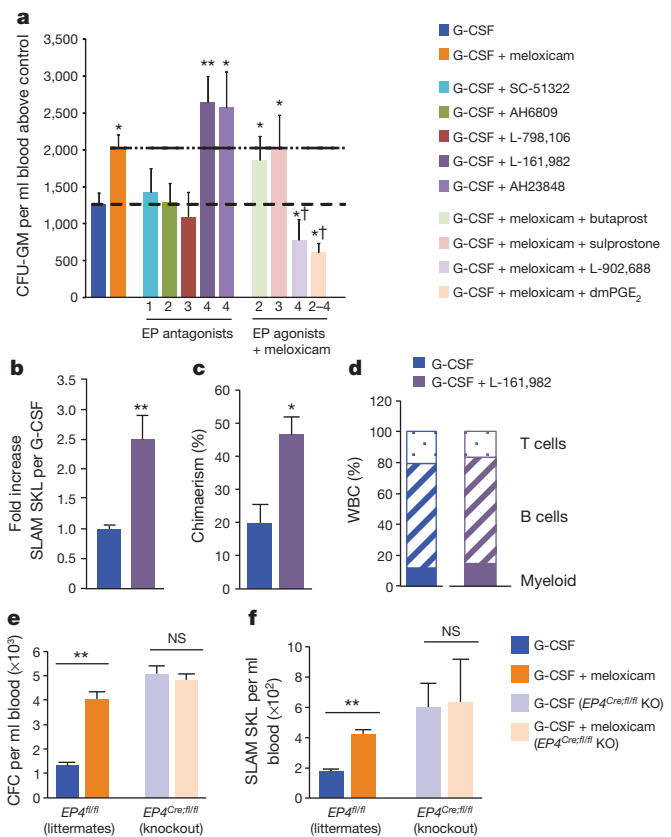


**Figure 2 | Non-human primates and healthy human volunteers mobilize HSCs and HPCs in response to NSAID treatment.** **a–e**, Four baboons were treated with G-CSF with/without meloxicam in a crossover design (shown in **a**). **b, c**, CD34<sup>+</sup> cells (**b**) and CFU-GM (**c**) in peripheral blood were determined. **d, e**, CD34<sup>+</sup> cells (**d**) and CFU-GM (**e**) in peripheral blood were determined before (pre) and after (post) five days of meloxicam treatment. **f–i**, Seven healthy human volunteers were treated with 15 mg meloxicam a day orally for five days, and assessed for CD34<sup>+</sup> cells (**f**), CFU-GM (**g**), BFU-E (**h**) and CFU-GEMM (**i**), before and after treatment. Statistics represent paired, two-tailed *t*-test; NS, not significant.

COX-1 and COX-2 and, therefore, there may be similar activity of NSAIDs with different COX-1 or COX-2 inhibitory profiles when compared to meloxicam.

Meloxicam, similar to indomethacin, increased the egress of HPCs (Fig. 1a and Supplementary Fig. 4) and the phenotypic HSC-enriched populations Sca-1<sup>+</sup> c-kit<sup>+</sup> Lin<sup>−</sup> (SKL) or the highly purified CD150<sup>+</sup> CD48<sup>−</sup> (SLAM) SKL populations (Fig. 1b and Supplementary Fig. 4). Enhanced cell egress was maintained in *Alox5* (also known as *5LX* and *5-ALOX*) knockout mice (Supplementary Fig. 5), further demonstrating that the effects are not due to general eicosanoid inhibition. Enhancement in cell egress was also not specific to G-CSF, as meloxicam enhanced the cell mobilization caused by the clinically used CXCR4 antagonist AMD3100 (Supplementary Fig. 6).

Despite significant increases in phenotypic HSCs and functional HPCs in the peripheral blood, two early transplant attempts did not show enhanced HSC engraftment (Supplementary Fig. 7a, b). Because we previously showed that PGE<sub>2</sub> signalling was a positive regulator of HSC CXCR4 expression and homing to the niche<sup>6</sup>, we proposed that although the HSC and HPC yield was increased in NSAID grafts, CXCR4 expression might be reduced, accounting for the apparent lack of enhanced engraftment. To test this hypothesis we staggered the administration of NSAID and G-CSF to allow for haematopoietic mobilization and restoration of normal endogenous PGE<sub>2</sub> signalling



**Figure 3 | PGE<sub>2</sub> EP4 receptor antagonism/knockout expands bone marrow HPCs and enhances mobilization.** **a**, HPC mobilization with G-CSF, G-CSF plus meloxicam, G-CSF plus E-prostanoid (EP) receptor antagonists, or G-CSF, meloxicam and E-prostanoid receptor agonists ( $n = 5$  mice per group, assayed individually). **b–d**, The EP4 antagonist L-161,982 enhanced HSC mobilization ( $n = 4$  mice per group, assayed individually) (**b**), and long-term reconstitution 16 weeks after transplantation (**c**), with multi-lineage reconstitution ( $n = 5$  mice per group, assayed individually) (**d**). WBC, white blood cells. **e, f**, EP4 knockout mobilization. Meloxicam enhances mobilization of HPCs (**e**) and SLAM SKL cells (**f**) in wild-type littermates (*EP4<sup>fl/fl</sup>*), but not in EP4 conditional knockouts (*EP4<sup>Cre/fl/fl</sup>* KO) ( $n = 3–4$  mice per group, assayed individually). \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; unpaired two-tailed *t*-test. † $P < 0.05$  compared to G-CSF plus meloxicam. All data are mean  $\pm$  s.e.m.

before transplantation (Supplementary Fig. 7c). CXCR4 levels were significantly lower after NSAID treatment, and staggered administration allowed for restored receptor levels, while maintaining enhanced HSC egress (Supplementary Fig. 7d, e). We competitively transplanted mobilized grafts from G-CSF, or non-staggered and staggered G-CSF plus meloxicam-treated mice. Staggered administration resulted in significant enhancement of long-term repopulating HSC (LT-HSC) engraftment, with a 48-h stagger resulting in a 2.6-fold increase in LT-HSCs (Fig. 1c–e and Supplementary Fig. 8). When grafts were transplanted non-competitively, staggered co-administration of meloxicam resulted in four-day faster recovery of neutrophils (Fig. 1f) and platelets (Fig. 1g) compared with G-CSF alone. Secondary transplantation confirmed sustained LT-HSC activity with multi-lineage reconstitution 36 weeks after transplant (Supplementary Fig. 9).

To confirm NSAID-mediated HSC and HPC egress in other species, four baboons were treated with a standard regimen of G-CSF, or with the combination of G-CSF plus meloxicam in a crossover design (Fig. 2a). Although the responses of individual baboons to G-CSF varied, in all cases meloxicam treatment increased CD34<sup>+</sup> cells (Fig. 2b) and granulocytic-monocytic progenitors (colony-forming unit-granulocyte, macrophage; CFU-GM) (Fig. 2c) in peripheral blood. Meloxicam treatment on its own also resulted in significant

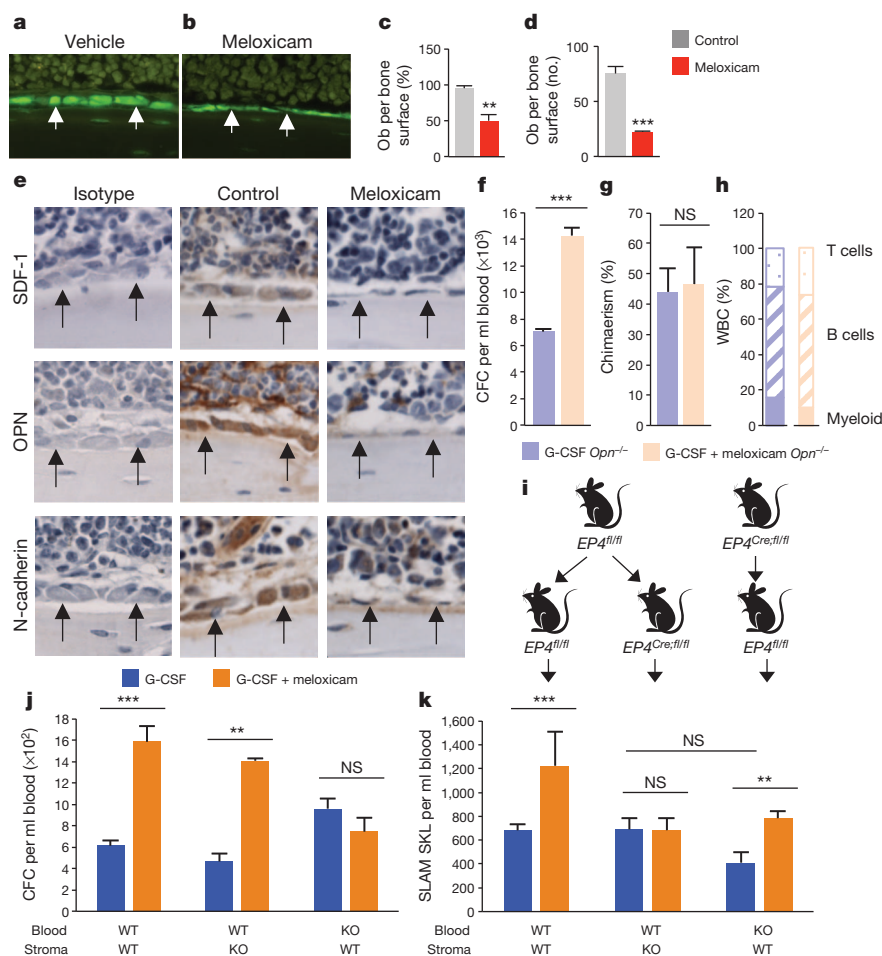


HSC and HPC egress (Fig. 2d, e). In healthy human volunteers, meloxicam treatment resulted in significant increases in CD34<sup>+</sup> cells (Fig. 2f), and functionally defined HPCs (Fig. 2g–i), matching the HSC and HPC egress seen with meloxicam treatment in baboons and mice. Thus, short-term endogenous PGE<sub>2</sub> inhibition, closely resembling current clinical NSAID treatment, results in a previously unappreciated increase in HSC and HPC mobilization.

Meloxicam treatment increased functionally defined myeloid progenitors and phenotypically defined granulocyte–macrophage progenitors in the bone marrow, but no differences in phenotypically or functionally defined HSC were observed (Supplementary Fig. 10). Because PGE<sub>2</sub> signals through four receptors (EP1–4), each with unique signalling pathways<sup>10</sup>, we proposed that the myeloid expansion and egress was due to lack of signalling via one or more of the E-prostanoid receptors. Only agonists capable of activating the EP4 receptor inhibited myeloid HPCs (Supplementary Fig. 11a). To confirm further the specific role of the EP4 receptor, similar assays were performed using knockout mice for each of the E-prostanoid receptors. Comparison of all knockout strains showed that only HPCs from conditional *EP4*<sup>-/-</sup> (also known as *Ptger4*<sup>-/-</sup>) mice had reduced responses to inhibition by PGE<sub>2</sub> (Supplementary Fig. 11b), and a 2.3-fold increase in marrow macrophage progenitors (colony-forming

unit-macrophage; CFU-M) compared to wild type (Supplementary Fig. 11c). Co-administration of EP4 antagonists with G-CSF significantly enhanced mobilization, similar to meloxicam, whereas EP1, EP2 and EP3 antagonists failed to increase mobilization (Fig. 3a). Furthermore, when a selective EP4 agonist was co-administered with G-CSF plus meloxicam, the meloxicam enhancement of mobilization was abrogated, and to the same degree as co-administration of the long-acting PGE<sub>2</sub> analogue 16,16-dimethyl-PGE<sub>2</sub> (dmPGE<sub>2</sub>). Agonists that did not target the EP4 receptor failed to alter meloxicam enhancement. EP4 antagonism with G-CSF enhanced mobilization of LT-HSCs (Fig. 3b–d), indicating that the NSAID-mediated effects on HPC/HSC egress are due to reduced EP4 receptor signalling. Consistent with pharmacological data, conditional *EP4* deletion increased HPC and HSC egress (Supplementary Fig. 11d–f), and enhanced mobilization caused by meloxicam was abrogated (Fig. 3e, f). These data implicate PGE<sub>2</sub> and EP4 receptor signalling in mediating the cell egress effects of NSAIDs; however, because we did not conduct a comprehensive lipidomic profile we cannot exclude contributions of other eicosanoids.

*In vitro* and *in vivo* results indicate that a lack of EP4 signalling drives HPC expansion, possibly explaining one mechanism responsible for the enhanced HPC egress: more marrow HPCs allows more to



**Figure 4** | NSAIDs attenuate haematopoietic supportive molecules and differentially mobilize HSCs and HPCs in *Opn* knockout and *EP4* conditional knockout mice. **a–d**, Assessment of Col2.3-GFP cells after vehicle (**a**) or meloxicam (**b**) demonstrates reduced percentages (**c**) and numbers (**d**) of osteolineage cells (expressed as the percentage/number of osteoblasts (ob) per bone surface) ( $n = 4$  mice per group, assayed individually). **e**, Immunohistochemical staining of haematopoietic supportive molecules after treatment with meloxicam (original magnification,  $\times 400$ ). Arrows denote osteolineage cell layer lining the endosteum. **f–h**, Meloxicam enhances mobilization of HPCs in *Opn*<sup>-/-</sup> mice

(**f**), with no enhancement in long-term reconstitution 16 weeks after transplantation (**g**, **h**). **i**, Representation of chimaera generation allowing conditional knockout of donor haematopoietic cells, or recipient stromal cells. *EP4* was deleted with tamoxifen 8 weeks after transplantation and mice were treated with G-CSF or G-CSF plus meloxicam. **j**, **k**, Enhanced mobilization of HPCs by meloxicam when *EP4* is expressed on haematopoietic cells (**j**) and enhanced mobilization of HSCs when *EP4* is expressed by stromal cells (**k**) ( $n = 4$  mice per group, assayed individually). WT, wild type. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; unpaired two-tailed *t*-test. Data are mean  $\pm$  s.e.m.

be mobilized to the periphery. However, no alterations in bone marrow HSC content were observed (Supplementary Fig. 10), suggesting that HSC mobilization results from a different mechanism, perhaps acting on the HSC niche. Gross histological analysis of NSAID-treated mice over 0–4 days showed a progressive increase in the laminarity of osteolineage cells lining the endosteum (Supplementary Figs 12 and 13), similar to that seen after G-CSF treatment<sup>11</sup>. Comparable results were observed in collagen 2.3-GFP reporter mice (which express the enhanced green fluorescent protein gene 2.3 kilobases upstream of the rat *Col1a1* promoter), showing marked attenuation of osteolineage cells (Fig. 4a–d), and in mice after conditional *EP4* deletion (Supplementary Fig. 14). Dynamic bone formation assays using staggered calcein double labelling and modified Goldner's trichrome staining support significant attenuation of osteolineage cellular function (Supplementary Fig. 15).

At present there is considerable debate about the direct or indirect roles of osteoclasts in haematopoietic niche regulation and HSC or HPC retention (reviewed in refs 12 and 13). To assess the role of osteoclasts, mice were treated with meloxicam and/or G-CSF with or without zoledronic acid, a potent inhibitor of osteoclast activity<sup>14</sup>. Similar to a recent report<sup>15</sup>, zoledronic acid resulted in an increase in HSC or HPC mobilization by meloxicam and G-CSF (Supplementary Fig. 16), suggesting that increased osteoclast activity is not a mitigating mechanism for NSAID-mediated HSC and HPC egress. Niche attenuation and HSC or HPC mobilization by G-CSF have recently been reported to be mediated by marrow-resident monocyte/macrophage populations<sup>15–17</sup>. In contrast to G-CSF<sup>15</sup>, immunohistochemical analysis demonstrated that meloxicam does not reduce F4/80<sup>+</sup> macrophages (Supplementary Fig. 17a), nor is there a reduction in phenotypically defined macrophages assessed by flow cytometry (Supplementary Fig. 17b, c). We observed no changes in sinusoidal endothelial cell number or apoptotic state (Supplementary Fig. 18), nor sinusoid vessels or endothelial cell number by immunohistochemistry (Supplementary Fig. 19). Similarly, there was no alteration in the nestin<sup>+</sup> cell number (Supplementary Fig. 20). No differences in marrow MMP-9 or soluble c-kit, agents reported to regulate HSC motility in the bone marrow niche<sup>18</sup>, were observed in NSAID-treated mice (data not shown), suggesting other unique HSC-retentive molecule(s) are regulated by *EP4*.

We fractionated osteolineage cells into three sub-populations<sup>19,20</sup> (Supplementary Fig. 21a). Quantitative PCR with reverse transcription (qRT-PCR) analysis revealed that all three populations expressed all four E-prostanoid receptors, with *EP4* being expressed most predominately (Supplementary Fig. 21b). Meloxicam treatment resulted in reductions in messenger RNA expression of several haematopoietic supportive molecules, including *Jag1*, *Runx2*, *Vcam1*, *SCF* (also known as *Kitl*), *SDF-1* (*Cxcl12*), and osteopontin (*Opn*, also known as *Spp1*) (Supplementary Fig. 21c). Similarly, immunohistochemistry staining demonstrated reductions in SDF-1, OPN and N-cadherin expression (Fig. 4e). Analysis in *EP4* conditional knockout mice showed a significant reduction in mesenchymal progenitor cells compared to Cre(–) littermates and wild-type controls (Supplementary Fig. 21d), further demonstrating a role for *EP4* signalling in haematopoietic niche maintenance.

Because the interaction of SDF-1 with its cognate receptor CXCR4 is a well-known mediator of niche retention, we sought to determine whether reduced expression of SDF-1 mediated the HSC and HPC egress caused by NSAID treatment. Surprisingly, despite the robust egress of cells in *Cxcr4* conditional knockout mice, both HPC and HSC trafficking to the periphery were significantly enhanced by meloxicam (Supplementary Fig. 22). Osteopontin has been reported as a regulator of both HSC quiescence<sup>21</sup> and niche retention<sup>22</sup>. In contrast to CXCR4, when *Opn* knockout mice were treated with meloxicam or G-CSF for six days, meloxicam enhanced the mobilization of HPCs (Fig. 4f), but, unexpectedly, not of HSCs (Fig. 4g, h; additional data in Supplementary Fig. 23), whereas both HPCs and HSCs were mobilized

by G-CSF in wild-type mice. This surprising result indicates that NSAID-mediated OPN reduction is specifically responsible for the observed HSC niche egress, whereas the increased number of peripheral HPCs results from an independent mechanism(s). To determine the differential roles of haematopoietic intrinsic versus stromal niche *EP4* signalling in mediating HPC/HSC egress, we created chimaeric mice in which we could conditionally delete *EP4* from donor haematopoietic cells or recipient stromal cells (Fig. 4i). *EP4* expression on haematopoietic cells was required for NSAID-mediated egress of HPCs (Fig. 4j), whereas *EP4* on stromal cells was specifically necessary for HSC egress (Fig. 4k). These studies demonstrate that PGE<sub>2</sub> signalling differentially regulates HPC and HSC retention in the bone marrow through both cell intrinsic and cell extrinsic mechanisms, and future studies should define the relative roles of individual stromal niche cell contributions to *EP4*-mediated niche retention. To our knowledge, this is the first report of an agent capable of mobilizing both HSCs and HPCs and doing so through cell-stage-specific mechanisms.

## METHODS SUMMARY

C57Bl/6 and *Opn*<sup>–/–</sup> mice were purchased from Jackson Laboratories. B6.SJL-PtrCAp3B/BoyJ mice were bred in-house. *Cxcr4*<sup>fllox/flox</sup> mice were generated as described<sup>23</sup> and were a gift from Y. R. Zou. *EPI*<sup>–/–</sup> (also known as *Ptger1*<sup>–/–</sup>), *EP2*<sup>–/–</sup> (*Ptger2*<sup>–/–</sup>), *EP3*<sup>–/–</sup> (*Ptger3*<sup>–/–</sup>), and *EP4*<sup>fllox/flox</sup> mice were generated as described<sup>24–26</sup>. Conditional mice were bred to *Ubc-Cre/ERT2* mice from Jackson. Female olive baboons (*Papio Anubis*) were housed individually in conventional caging of the Biological Resources Laboratory, University of Illinois at Chicago. Primate research was approved by the University of Illinois Animal Care and Use Committee (IACUC). The IACUC of the Indiana University School of Medicine (IUSM) approved all protocols. The Institutional Review Board of IUSM approved human subject research and informed consent was acquired from all volunteers.

**Full Methods** and any associated references are available in the online version of the paper.

Received 6 May 2012; accepted 22 January 2013.

Published online 13 March 2013.

- Calvi, L. M. *et al.* Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature* **425**, 841–846 (2003).
- Ding, L., Saunders, T. L., Enikolopov, G. & Morrison, S. J. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature* **481**, 457–462 (2012).
- Méndez-Ferrer, S. *et al.* Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature* **466**, 829–834 (2010).
- Raaijmakers, M. H. *et al.* Bone progenitor dysfunction induces myelodysplasia and secondary leukaemia. *Nature* **464**, 852–857 (2010).
- Zhang, J. *et al.* Identification of the haematopoietic stem cell niche and control of the niche size. *Nature* **425**, 836–841 (2003).
- Hoggatt, J., Singh, P., Sampath, J. & Pelus, L. M. Prostaglandin E<sub>2</sub> enhances hematopoietic stem cell homing, survival, and proliferation. *Blood* **113**, 5444–5455 (2009).
- North, T. E. *et al.* Prostaglandin E<sub>2</sub> regulates vertebrate haematopoietic stem cell homeostasis. *Nature* **447**, 1007–1011 (2007).
- Ahmed, M., Khanna, D. & Furst, D. E. Meloxicam in rheumatoid arthritis. *Expert Opin. Drug Metab. Toxicol.* **1**, 739–751 (2005).
- Rinder, H. M. *et al.* Effects of meloxicam on platelet function in healthy adults: a randomized, double-blind, placebo-controlled trial. *J. Clin. Pharmacol.* **42**, 881–886 (2002).
- Breyer, R. M., Bagdassarian, C. K., Myers, S. A. & Breyer, M. D. Prostanoid receptors: subtypes and signaling. *Annu. Rev. Pharmacol. Toxicol.* **41**, 661–690 (2001).
- Katayama, Y. *et al.* Signals from the sympathetic nervous system regulate hematopoietic stem cell egress from bone marrow. *Cell* **124**, 407–421 (2006).
- Bethel, M., Srour, E. F. & Kacena, M. A. Hematopoietic cell regulation of osteoblast proliferation and differentiation. *Curr. Osteoporos. Rep.* **9**, 96–102 (2011).
- Hoggatt, J. & Pelus, L. M. Many mechanisms mediating mobilization: an alliterative review. *Curr. Opin. Hematol.* **18**, 231–238 (2011).
- Mundy, G. R., Yoneda, T. & Hiraga, T. Preclinical studies with zoledronic acid and other bisphosphonates: impact on the bone microenvironment. *Semin. Oncol.* **28**, 35–44 (2001).
- Winkler, I. G. *et al.* Bone marrow macrophages maintain hematopoietic stem cell (HSC) niches and their depletion mobilizes HSCs. *Blood* **116**, 4815–4828 (2010).
- Chow, A. *et al.* Bone marrow CD169<sup>+</sup> macrophages promote the retention of hematopoietic stem and progenitor cells in the mesenchymal stem cell niche. *J. Exp. Med.* **208**, 261–271 (2011).
- Christopher, M. J., Rao, M., Liu, F., Woloszynek, J. R. & Link, D. C. Expression of the G-CSF receptor in monocytic cells is sufficient to mediate hematopoietic progenitor mobilization by G-CSF in mice. *J. Exp. Med.* **208**, 251–260 (2011).
- Heissig, B. *et al.* Recruitment of stem and progenitor cells from the bone marrow niche requires MMP-9 mediated release of Kit-ligand. *Cell* **109**, 625–637 (2002).

19. Chitteti, B. R. *et al.* Impact of interactions of cellular components of the bone marrow microenvironment on hematopoietic stem and progenitor cell function. *Blood* **115**, 3239–3248 (2010).
20. Nakamura, Y. *et al.* Isolation and characterization of endosteal niche cell populations that regulate hematopoietic stem cells. *Blood* **116**, 1422–1432 (2010).
21. Stier, S. *et al.* Osteopontin is a hematopoietic stem cell niche component that negatively regulates stem cell pool size. *J. Exp. Med.* **201**, 1781–1791 (2005).
22. Grassinger, J. *et al.* Thrombin-cleaved osteopontin regulates hemopoietic stem and progenitor cell functions through interactions with  $\alpha_9\beta_1$  and  $\alpha_4\beta_1$  integrins. *Blood* **114**, 49–59 (2009).
23. Nie, Y. *et al.* The role of CXCR4 in maintaining peripheral B cell compartments and humoral immunity. *J. Exp. Med.* **200**, 1145–1156 (2004).
24. Kennedy, C. R. *et al.* Salt-sensitive hypertension and reduced fertility in mice lacking the prostaglandin EP<sub>2</sub> receptor. *Nature Med.* **5**, 217–220 (1999).
25. Guan, Y. *et al.* Antihypertensive effects of selective prostaglandin E<sub>2</sub> receptor subtype 1 targeting. *J. Clin. Invest.* **117**, 2496–2505 (2007).
26. Schneider, A. *et al.* Generation of a conditional allele of the mouse prostaglandin EP4 receptor. *Genesis* **40**, 7–14 (2004).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** These studies were supported by National Institutes of Health (NIH) grants HL096305 (L.M.P.), CA143057, CA069158 (T.A.G., K.S.M.), HL100402 (D.T.S.) and DK37097 (R.M.B.). J.H. was supported by NIH training grants DK07519, HL07910 and HL087735. Flow cytometry was performed in the Flow Cytometry Resource Facility of the Indiana University Simon Cancer Center (NCIP30 CA082709). Additional core support was provided by a Center of Excellence in Hematology grant P01 DK090948. The authors would like to thank H. E. Broxmeyer and B. Saez for critically reading the manuscript.

**Author Contributions** All authors assisted in writing of the manuscript. J.H. analysed data, wrote the manuscript, designed all experiments and implemented all experiments with assistance from P.S., A.F.H., B.R.C., J.M.S., P.H., B.A.P., K.N.S., F.F., L.S. and F.K.W. K.S.M., M.C. and T.A.G. performed histological analyses and assisted with corresponding study designs. G.L.M. and R.M.B. performed eicosanoid analysis and generated E-prostanoid receptor knockout mice, and C.H.S. assisted with *Alox5* mice and experiments. D.T.S. and E.F.S. assisted with experimental design and data analyses. L.M.P. designed and performed experiments, analysed and evaluated all data, and wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.M.P. ([lpelus@iupui.edu](mailto:lpelus@iupui.edu)).



## METHODS

**Animals and subjects.** C57Bl/6 (CD45.2) mice were purchased from Jackson Laboratories (Bar Harbour). B6.SJL-PtrcAPep3B/BoyJ (BOYJ) (CD45.1) mice were bred in-house. *Cxcr4<sup>fllox/fllox</sup>* mice were generated as described<sup>23</sup> and were a gift from Y.-R. Zou. *EP1<sup>-/-</sup>* (also known as *Ptger1<sup>-/-</sup>*), *EP2<sup>-/-</sup>* (*Ptger2<sup>-/-</sup>*), *EP3<sup>-/-</sup>* (*Ptger3<sup>-/-</sup>*), and *EP4<sup>fllox/fllox</sup>* mice were generated as described<sup>24–26</sup>. *Opn<sup>-/-</sup>* mice were purchased from Jackson Laboratories. *Nestin-GFP<sup>27</sup>*, *Col2.3-GFP<sup>28</sup>* and *Alox5* (ref. 29) mice were generated as described. *EP4<sup>fllox/fllox</sup>* mice were bred to *Ubc-Cre/ERT2* mice from Jackson Laboratories to generate conditional *EP4* knockout mice. All mice were maintained on a C57Bl/6 background. Female olive baboons (*Papio Anubis*), within the weight range of 16–19 kg, were housed individually in conventional caging and holding rooms of the Biological Resources Laboratory, a centralized animal facility for the University of Illinois at Chicago Medical Center. The conducted primate research was approved by the University of Illinois at Chicago Animal Care and Use Committee. The Animal Care and Use Committee of Indiana University School of Medicine (IUSM) approved all protocols, and the Institutional Review Board approved human subject research. Informed consent was obtained from all volunteers.

**Peripheral blood and bone marrow acquisition and processing.** Peripheral blood from mice was obtained by cardiac puncture after CO<sub>2</sub> asphyxiation using an EDTA-rinsed syringe. Blood was transferred to tubes containing EDTA for complete blood cell (CBC) analysis. CBC analysis was performed on a Hemavet 950FS (Drew Scientific). PBMCs were prepared by centrifugation over lympholyte mammal (Cedarlane Laboratories) at 800 g for 30–40 min at room temperature, followed by triplicate washes. Bone marrow cells were collected by flushing femurs with ice-cold PBS, and single-cell suspensions were prepared by passage through a 26-gauge needle. For baboons, peripheral blood was obtained from the femoral vein of baboons anaesthetized with an intramuscular injection of 10 mg kg<sup>-1</sup> ketamine hydrochloride (Bionichepharma). Blood was collected into 10-ml sterile EDTA vacutainers (Becton Dickinson) and transported on ice to IUSM for analysis. Complete blood counts with differentials were performed on a Hemavet 950FS. Peripheral blood was then diluted 1:3 with PBS and mononuclear cells were isolated using Ficoll-Paque Plus (Amersham Biosciences), per manufacturer's protocol.

**Colony assays.** Bone marrow cells or PBMCs were resuspended in McCoy's 5A modified media supplemented with 100 U ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin, 0.6× modified essential medium (MEM) vitamin solution, 1 mM sodium pyruvate, 0.8× MEM essential amino acids, 0.6× MEM non-essential amino acids, 0.05% sodium bicarbonate (all from Gibco), serine, asparagine, glutamine mixture and 15% HI-FBS (Hyclone Sterile Systems) as described<sup>30,31</sup>. Cells were mixed with 0.3% agar (Difco Laboratories) in McCoy's 5A medium with 10 ng ml<sup>-1</sup> rhGM-CSF and 50 ng ml<sup>-1</sup> rmSCF (R&D Systems). PBMCs were cultured at 2 × 10<sup>5</sup> cells per ml and bone marrow cells at 5 × 10<sup>4</sup> cells per millilitre. All cultures were established in triplicate from individual animals, incubated at 37 °C, 5% CO<sub>2</sub>, 5% O<sub>2</sub> in air for seven days and colonies were quantified by microscopy. In some experiments, total CFC including CFU-GM, BFU-E and CFU-GEMM were enumerated in 1% methylcellulose/IMDM media containing 30% FBS, 1 U ml<sup>-1</sup> recombinant human erythropoietin (EPO), 10 ng ml<sup>-1</sup> rhGM-CSF or rmGM-CSF and 50 ng ml<sup>-1</sup> rhSCF or rmSCF as described<sup>32,33</sup>. In some experiments, phenotypically defined common myeloid progenitor (CMP) and granulocyte-macrophage progenitor (GMP) were plated at 500 cells per plate and colony growth determined in agar CFC assays with rmGM-CSF plus rmSCF or rmM-CSF. For analysis of CFC in baboons, similar assays were performed using recombinant human growth factors.

**Flow cytometry.** All antibodies were purchased from BD Biosciences unless otherwise noted. For detection of SKL cells, we used streptavidin conjugated with phycoerythrin (PE)-Cy7 (to stain for biotinylated MACS lineage antibodies (Miltenyi)), c-kit-allophycocyanin (APC), Sca-1-PE or APC-Cy7, CD45.1-PE or CD45.2-FITC. For SLAM SKL, we used Sca-1-PE-Cy7, c-kit-FITC, CD150-APC (eBiosciences), CD48-biotin (eBiosciences) and streptavidin-PE. CXCR4 expression was analysed using biotinylated lineage antibodies, streptavidin-PE-Cy7, c-kit-APC, Sca-1-APC-Cy7 and CXCR4-PE. For baboon CD34 analysis, CD34-PE (clone 563) was used. For macrophages, antibodies against CD115 (clone AFS98), Gr-1 (clone RB6-8C5) and F4/80 (clone CI:A3-1) were used. Osteolineage populations were identified and sorted as previously described<sup>19</sup>. For enumeration of bone marrow endothelial cells, femurs and tibias were crushed in a sterile mortar, and digested in collagenase (0.3%) at 37 °C for 1 h. Recovered cells were co-stained with fluorochrome-conjugated antibodies against CD45, Ter119, Sca-1, VEGFR3 and CD31, and the total number of sinusoidal endothelial cells (SECs; CD45<sup>+</sup> Ter119<sup>+</sup> Sca-1<sup>+</sup> VEGFR3<sup>+</sup> CD31<sup>+</sup>) per femur was enumerated by flow cytometry analysis. To examine endothelial cell apoptosis, gated CD45<sup>+</sup> Ter119<sup>+</sup> Sca-1<sup>+</sup> VEGFR3<sup>+</sup> CD31<sup>+</sup> cells were stained with annexin V (BD Biosciences) and LIVE/DEAD staining dye (Invitrogen). For enumeration of

myeloid progenitors (CMP, GMP and megakaryocyte-erythroid progenitors (MEP)), femurs and tibias were flushed with 5 ml IMDM containing 2% FBS. Lineage-positive cells were depleted using the lineage-cell depletion kit (Miltenyi Biotec), and lineage-negative cells were stained with fluorochrome-conjugated antibodies against Sca-1, c-Kit, IL-7R $\alpha$ , CD34 and FCRIII/III and analysed by flow cytometry. The Lin<sup>+</sup> IL-7R $\alpha$ <sup>+</sup> Sca-1<sup>+</sup> c-Kit<sup>+</sup> fraction was subdivided into three subpopulations, CMP (FCRIII/III<sup>low</sup> CD34<sup>+</sup>), MEP (FCRIII/III<sup>low</sup> CD34<sup>+</sup>), and GMP (FCRIII/III<sup>low</sup> CD34<sup>+</sup>), and collected by sorting. All flow cytometry analyses were performed on an LSRII flow cytometer (BD). Cell sorting was performed on a BD Aria or Reflection II or Reflection III sorters.

**Peripheral blood mobilization.** Several different mobilization strategies were used, with specific details of dosing and schematics of dosing regimens shown on the data figures or included in the figure legends. In general, mice were given subcutaneous treatments of vehicle, NSAIDs (at varying doses), G-CSF (50 µg kg<sup>-1</sup>, twice a day for four days), or G-CSF plus NSAIDs. For studies exploring mobilizing agents other than G-CSF, mice were treated with AMD3100 (5 mg kg<sup>-1</sup> at day five; single injection), and peripheral blood was collected 1 h after AMD3100 treatment. For comparisons of several different NSAIDs, all NSAIDs were dosed by oral gavage using an enhanced oral gavage technique<sup>34</sup>. Each gavage treatment was given in a 0.2-ml bolus (10 ml kg<sup>-1</sup>) of 0.5% methyl cellulose (Sigma-Aldrich) with an NSAID suspended in solution. For E-prostanoid receptor analysis, mice were mobilized with G-CSF in combination with meloxicam, AH6809 (EP1–3 antagonist, 10 µg per mouse, intraperitoneally, four days), AH23848 (EP4 antagonist, 10 µg per mouse, intraperitoneally, four days), L-161,982 (EP4 antagonist, 10 µg per mouse, intraperitoneally, four days) or G-CSF plus meloxicam and an EP2, EP1/3 or EP4 agonist (10 µg per mouse, intraperitoneally, four days) or dmPGE<sub>2</sub> (10 µg per mouse, intraperitoneally, four days). For baboon studies, a baseline bleed was performed for CBC, CD34 and CFC analysis. Two days later, two baboons were treated with 10 µg kg<sup>-1</sup> G-CSF, and two baboons were treated subcutaneously with 10 µg kg<sup>-1</sup> G-CSF and 0.2 mg kg<sup>-1</sup> meloxicam on day 1, followed by 0.1 mg kg<sup>-1</sup> meloxicam for subsequent days, for a total of five days. Blood was collected after the treatment regimen for CBC, CD34 and CFC analysis. After a two-week resting period, the above procedure was repeated, switching treatment groups for individual baboons. Furthermore, after another two-week resting period, blood was collected before and after a five-day treatment regimen with meloxicam, and CBC, CD34 and CFC were analysed. For healthy volunteer studies, subjects naive to any medications within 30-days received a baseline bleed, followed by a second bleed after a five-day regimen of 15 mg meloxicam per day, orally. CD34 cells were assessed by the International Society of Hematology and Graft Engineering (ISHAGE) procedure<sup>35</sup> performed by the Stem Cell Laboratory of the IUSM Bone Marrow Transplant Program. CFC were assessed as described above.

**Limiting dilution competitive transplantation.** CD45.1 mice were mobilized with a standard four-day regimen of G-CSF, or G-CSF plus a four-day regimen of meloxicam (6 mg kg<sup>-1</sup>). In some studies designed to evaluate timing and duration of NSAID dosing in combination with G-CSF, initiation of the NSAID regimen preceded G-CSF treatment and was staggered such that NSAID administration ended simultaneous with the G-CSF regimen (no stagger), one day before G-CSF (1-day stagger) or two days before G-CSF (2-day stagger) (regimens as depicted in the corresponding data figure). On day 5, PBMCs were acquired and transplanted at 1:1, 2:1, 3:1 or 4:1 ratios with 5 × 10<sup>5</sup> C57Bl/6J WBM competitors into lethally irradiated C57Bl/6J recipient mice. Peripheral blood chimaerism was monitored monthly, and CRU and LT-HSC frequency calculated. Transplants to evaluate LT-HSC mobilized in *Opn<sup>-/-</sup>* mice or with EP4 antagonist were performed competitively at a 4:1 ratio; 800,000 PBMCs from CD45.2 mice versus 200,000 WBM from CD45.1 mice and peripheral blood chimaerism and multilineage reconstitution were assessed 16 weeks after transplantation.

**Recovery assay.** Mice were mobilized with G-CSF or G-CSF plus meloxicam with staggered dosing as described above, and 2 × 10<sup>6</sup> mobilized PBMCs were transplanted non-competitively into cohorts of ten lethally irradiated recipients per group. A cohort of non-irradiated mice was bled on the same schedule as the experimental-treated groups of mice. Every other day, five mice from each group were bled (~50 µl from a tail snip), and neutrophils and platelets in blood were enumerated using a Hemavet 950FS. Alternative groups of five mice were bled on each successive bleeding time point so that mice were only bled once every four days. Recovery of neutrophils and platelets to 50% and 100% were determined by comparison to the average neutrophil and platelet counts in the control group throughout the experimental period. After 90 days, mice were euthanized, bone marrow was collected and transplanted at a 2.5:1 ratio with 2 × 10<sup>5</sup> congenic competitors into lethally irradiated recipients to determine long-term repopulating ability of the primary mobilized graft.

**EP4 chimaera generation and mobilization assay.** Chimeras were generated using *EP4<sup>Cre/fllox/fllox</sup>* mice and age- and sex-matched *EP4<sup>fllox/fllox</sup>* littermate controls. *EP4<sup>fllox/fllox</sup>* mice were lethally irradiated and transplanted with 2 × 10<sup>6</sup> WBM cells from either

$EP4^{flox/flox}$  mice, allowing for generation of a wild-type:wild-type chimaera, or from  $EP4^{Cre/flox/flox}$  mice, allowing for generation of a knockout:wild-type chimaera. Similarly,  $EP4^{Cre/flox/flox}$  mice were lethally irradiated and transplanted with  $2 \times 10^6$  WBM cells from  $EP4^{flox/flox}$  mice, allowing for generation of a wild-type:knockout chimaera. At 8 weeks after transplantation, all mice were treated with 2 mg tamoxifen for three consecutive days, rested for three days and injected for three more days. Mice were then treated with G-CSF or G-CSF plus meloxicam starting ten days after the last treatment, and peripheral blood CFC and SLAM SKL were assessed as described.  $EP4$  gene deletion was confirmed by qRT-PCR.

**qRT-PCR.** For E-prostanoid receptor expression on sorted osteolineage cells, total RNA was extracted with the purelink RNA micro kit (Invitrogen). On-column DNase treatment was performed according to the manufacturer's instructions to eliminate contaminating genomic DNA. Conventional reverse transcription was followed with SuperScript III first-strand synthesis system (Invitrogen). qRT-PCR was performed using the SYBR advantage qPCR Premix kit (Clontech) on MxPro-3000 (Agilent). Primers were synthesized at IDT (Supplementary Table 1). A primer concentration of 250 nM was found to be optimal in all cases. The PCR protocol consisted of one cycle at 95 °C (5 min) followed by 45 cycles of 95 °C (15 s), 55 °C (30 s) and 72 °C (30 s). The dissociation curves were determined on each analysis to confirm that only one product was obtained. Expression of glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) and hypoxanthine guanine phosphoribosyl transferase (*HPRT*) were used as reference genes. The average threshold cycle number ( $C_t$ ) for each tested mRNA was used to quantify the relative expression of each gene. For analysis of haematopoietic supportive molecules on sorted osteolineage cells from vehicle-treated or NSAID-treated mice, qRT-PCR was performed with the TaqMan gene expression assay kit (Life Technologies) (Supplementary Table 2), with complementary DNA generated from the high capacity cDNA reverse transcription kit (Life Technologies). Microfluidic qRT-PCR was performed on BioMark Dynamic Arrays according to manufacturer's instructions (Fluidigm Corporation).

**Micro-computed tomography.** Formalin-fixed tibiae and femora were imaged with micro-computed tomography using a microCT-viva 40 (Scanco Medical AG) with a voxel size of 10.5 µm in all dimensions ( $n = 5$ ). The bones were mounted in a cylindrical specimen holder to be captured in a single scan. Bones were secured in the specimen holder with gauze and were completely submerged in 70% ethanol. The region of interest comprised 100 transverse computed tomography slices. Scans with an isotropic resolution of 10.5 µm were made using a 55-kV peak voltage X-ray beam. The fractional trabecular bone volume (BV/TV) and architectural properties of trabecular reconstructions, apparent trabecular thickness (Tb.Th.), trabecular number (Tb.N.), trabecular spacing (Tb.Sp.), and connectivity density (Conn.D.) were calculated.

**Dynamic and static bone histomorphometry.** Dynamic bone formation assays were performed using staggered calcein double labelling, as described previously<sup>36</sup>. Bone histomorphometry was performed on 7-µm thick sections of undecalcified femurs embedded in methylmethacrylate using standard procedures. The mineral apposition rate (mm per day), mineralizing surface and bone formation rate ( $\text{mm}^3 \text{mm}^{-2}$  per day) were measured on femora. Modified Goldner's trichrome staining procedure was performed on 7-µm thick sections of undecalcified femurs embedded in methylmethacrylate. The osteoid surfaces as well as quiescent surfaces were measured on the tissue sections. Bone marrow sinusoids were visualized with anti-VEGFR3 on 3.5-µm section. Vessels were identified by the positive staining around the vessel walls, and vessel areas were measured using the automated measuring system and expressed as a percentage or tissue volume. Vessel surface was traced with the same automated system. Vessel wall that showed an intact epithelial surface was expressed as endothelial surface over total vessel surface. For Col2.3-GFP analysis, 3.5-µm thick sections were obtained from treated Col2.3-GFP mice. Sections were visualized with a fluorescent microscope (Leica D100) using a FITC filter. Images were captured at  $\times 400$  magnification at four different areas in the mid-shaft of the femur. GFP<sup>+</sup> osteoblasts were counted on endocortical bone surfaces, and data were expressed as the number of osteoblasts per endocortical bone surface. Osteoblast surface was measured and expressed as

the endocortical bone surface covered by osteoblasts. All histomorphometry experiments were done on images captured using a Leica microscope outfitted with Q-imaging camera (W. Nuhlsbaum), and the histomorphometry analysis was done using Bioquant Osteo software automated measuring system (Bioquant Imaging Corporation). All histomorphometry values were expressed according to the standard nomenclature<sup>37,38</sup>.

**Immunohistochemistry.** Immunohistochemical analysis was performed on decalcified paraffin-embedded tissue sections. Rabbit polyclonal antibodies against N-cadherin and SDF1 N-terminal were from Abcam. Secondary anti-rabbit antibodies were from Vector Laboratories. Anti-mouse OPN was from R&D Systems. Secondary biotinylated anti-goat horseradish peroxidase (HRP) conjugate, HRP-DAB system and DAV chromogen were from R&D Systems. HRP-rat IgG2 isotype was used as a primary antibody negative control for SDF-1, OPN and N-cadherin, at 1:50. Isotype staining control was performed under the same conditions as the antibody staining.

**COX metabolite and activity analysis.** Mice were subcutaneously treated with vehicle control or meloxicam, twice daily. One hour after the last treatment, femurs were pulled and flushed with 1 ml of ice-cold PBS, quickly brought to single-cell suspension and then flash frozen. COX-1- and COX-2-derived metabolites were assessed by gas chromatography-mass spectrometry as we have previously described<sup>39,40</sup>. The second femur was processed in an identical way, and COX-1 and COX-2 activity were determined using a fluorescent COX activity assay following the manufacturer's instructions (Cayman Chemicals).

27. Mignone, J. L., Kukekov, V., Chiang, A. S., Steindler, D. & Enikolopov, G. Neural stem and progenitor cells in nestin-GFP transgenic mice. *J. Comp. Neurol.* **469**, 311–324 (2004).
28. Kalajic, Z. *et al.* Directing the expression of a green fluorescent protein transgene in differentiated osteoblasts: comparison between rat type I collagen and rat osteocalcin promoters. *Bone* **31**, 654–660 (2002).
29. Chen, X. S., Sheller, J. R., Johnson, E. N. & Funk, C. D. Role of leukotrienes revealed by targeted disruption of the 5-lipoxygenase gene. *Nature* **372**, 179–182 (1994).
30. King, A. G. *et al.* Rapid mobilization of murine hematopoietic stem cells with enhanced engraftment properties and evaluation of hematopoietic progenitor cell mobilization in rhesus monkeys by a single injection of SB-251353, a specific truncated form of the human CXCL12 chemokine GRO $\beta$ . *Blood* **97**, 1534–1542 (2001).
31. Pelus, L. M., Broxmeyer, H. E., Kurland, J. I. & Moore, M. A. Regulation of macrophage and granulocyte proliferation. Specificities of prostaglandin E and lactoferrin. *J. Exp. Med.* **150**, 277–292 (1979).
32. Broxmeyer, H. E. *et al.* SDF-1/CXCL12 enhances in vitro replating capacity of murine and human multipotential and macrophage progenitor cells. *Stem Cells Dev.* **16**, 589–596 (2007).
33. Fukuda, S., Bian, H., King, A. G. & Pelus, L. M. The chemokine GRO $\beta$  mobilizes early hematopoietic stem cells characterized by enhanced homing and engraftment. *Blood* **110**, 860–869 (2007).
34. Hoggatt, A. F., Hoggatt, J., Honerlaw, M. & Pelus, L. M. A spoonful of sugar helps the medicine go down: a novel technique to improve oral gavage in mice. *J. Am. Assoc. Lab. Anim. Sci.* **49**, 329–334 (2010).
35. Sutherland, D. R., Anderson, L., Keeney, M., Nayar, R. & Chin-Yee, I. The ISHAGE guidelines for CD34<sup>+</sup> cell determination by flow cytometry. *J. Hematother.* **5**, 213–226 (1996).
36. Mohammad, K. S. *et al.* Pharmacologic inhibition of the TGF- $\beta$  type I receptor kinase has anabolic and anti-catabolic effects on bone. *PLoS ONE* **4**, e5275 (2009).
37. Rowe, P. S. *et al.* Correction of the mineralization defect in hyp mice treated with protease inhibitors CA074 and pepstatin. *Bone* **39**, 773–786 (2006).
38. Parfitt, A. M. *et al.* Bone histomorphometry: standardization of nomenclature, symbols, and units. Report of the ASBMR Histomorphometry Nomenclature Committee. *J. Bone Miner. Res.* **2**, 595–610 (1987).
39. Murali, G. *et al.* Fish oil and indomethacin in combination potentially reduce dyslipidemia and hepatic steatosis in *LDLR*<sup>-/-</sup> mice. *J. Lipid Res.* **53**, 2186–2197 (2012).
40. Liu, T. *et al.* Prostaglandin E<sub>2</sub> deficiency uncovers a dominant role for thromboxane A<sub>2</sub> in house dust mite-induced allergic pulmonary inflammation. *Proc. Natl. Acad. Sci. USA* **109**, 12692–12697 (2012).

# NANOG-dependent function of TET1 and TET2 in establishment of pluripotency

Yael Costa<sup>1\*</sup>, Junjun Ding<sup>2\*</sup>, Thorold W. Theunissen<sup>1,3\*</sup>, Francesco Faiola<sup>2\*</sup>, Timothy A. Hore<sup>4</sup>, Pavel V. Shliha<sup>5</sup>, Miguel Fidalgo<sup>2</sup>, Arven Saunders<sup>2</sup>, Moyra Lawrence<sup>1,3</sup>, Sabine Dietmann<sup>1</sup>, Satyabrata Das<sup>6</sup>, Dana N. Levasseur<sup>6</sup>, Zhe Li<sup>7</sup>, Mingjiang Xu<sup>7</sup>, Wolf Reik<sup>4,8</sup>, José C. R. Silva<sup>1,3\*</sup> & Jianlong Wang<sup>2\*</sup>

Molecular control of the pluripotent state is thought to reside in a core circuitry of master transcription factors including the homeodomain-containing protein NANOG<sup>1,2</sup>, which has an essential role in establishing ground state pluripotency during somatic cell reprogramming<sup>3,4</sup>. Whereas the genomic occupancy of NANOG has been extensively investigated, comparatively little is known about NANOG-associated proteins<sup>5</sup> and their contribution to the NANOG-mediated reprogramming process. Using enhanced purification techniques and a stringent computational algorithm, we identify 27 high-confidence protein interaction partners of NANOG in mouse embryonic stem cells. These consist of 19 previously unknown partners of NANOG that have not been reported before, including the ten-eleven translocation (TET) family methylcytosine hydroxylase TET1. We confirm physical association of NANOG with TET1, and demonstrate that TET1, in synergy with NANOG, enhances the efficiency of reprogramming. We also find physical association and reprogramming synergy of TET2 with NANOG, and demonstrate that knockdown of TET2 abolishes the reprogramming synergy of NANOG with a catalytically deficient mutant of TET1. These results indicate that the physical interaction between NANOG and TET1/TET2 proteins facilitates reprogramming in a manner that is dependent on the catalytic activity of TET1/TET2. TET1 and NANOG co-occupy genomic loci of genes associated with both maintenance of pluripotency and lineage commitment in embryonic stem cells, and TET1 binding is reduced upon NANOG depletion. Co-expression of NANOG and TET1 increases 5-hydroxymethylcytosine levels at the top-ranked common target loci *Esrrb* and *Oct4* (also called *Pou5f1*), resulting in priming of their expression before reprogramming to naive pluripotency. We propose that TET1 is recruited by NANOG to enhance the expression of a subset of key reprogramming target genes. These results provide an insight into the reprogramming mechanism of NANOG and uncover a new role for 5-methylcytosine hydroxylases in the establishment of naive pluripotency.

We expanded the NANOG interactome in mouse embryonic stem cells using an improved affinity purification and mass spectrometry (AP-MS) strategy<sup>6,7</sup> similar to that described previously<sup>8</sup>, combined with an interactomics analysis (see Supplementary Information). This analysis identified 27 high-confidence interaction partners of NANOG (Fig. 1a, b, Supplementary Figs 1–4 and Supplementary Tables 1 and 2). Notable among the 19 previously unknown interaction partners of NANOG was the methylcytosine hydroxylase TET1<sup>9,10</sup> (Fig. 1b). Specific association of TET1 with NANOG was detected in all five affinity purification runs of three independent affinity purifications (Supplementary Fig. 5b and Supplementary Table 2), and the interaction between NANOG and TET1 was further confirmed by immunoprecipitation and co-immunoprecipitation (Fig. 1c and Supplementary

Fig. 5c, d). Whereas NANOG clearly associates with TET1 in embryonic stem cells, there also exists TET1-free NANOG protein as shown by immunodepleting TET1 in embryonic stem cells (Supplementary Fig. 5e). Notably, among the 27 high-confidence interaction partners of NANOG, at least 5 (NACC1, SGOL2, QSER1, HDAC2 and OCT4) were also associated with TET1 by co-immunoprecipitation and/or immunoprecipitation-mass spectrometry experiments (Supplementary Fig. 5f–h). Expression of TET1, like that of NANOG, is upregulated during reprogramming to pluripotency (Supplementary Fig. 6a). Because NANOG is a critical determinant during establishment of pluripotency<sup>3,11</sup>, we investigated whether TET1 may also be required for efficient nuclear reprogramming. Indeed, RNA interference-mediated inhibition of TET1 during reprogramming reduced generation of induced pluripotent stem (iPS) cells from mouse embryonic fibroblasts (Supplementary Fig. 6b–g and Supplementary Fig. 7). The requirement of TET1 for efficient reprogramming was confirmed using *Tet1*<sup>−/−</sup> embryonic stem cells<sup>12</sup> in an independent, heterokaryon-based reprogramming system<sup>13</sup> (Supplementary Fig. 8).

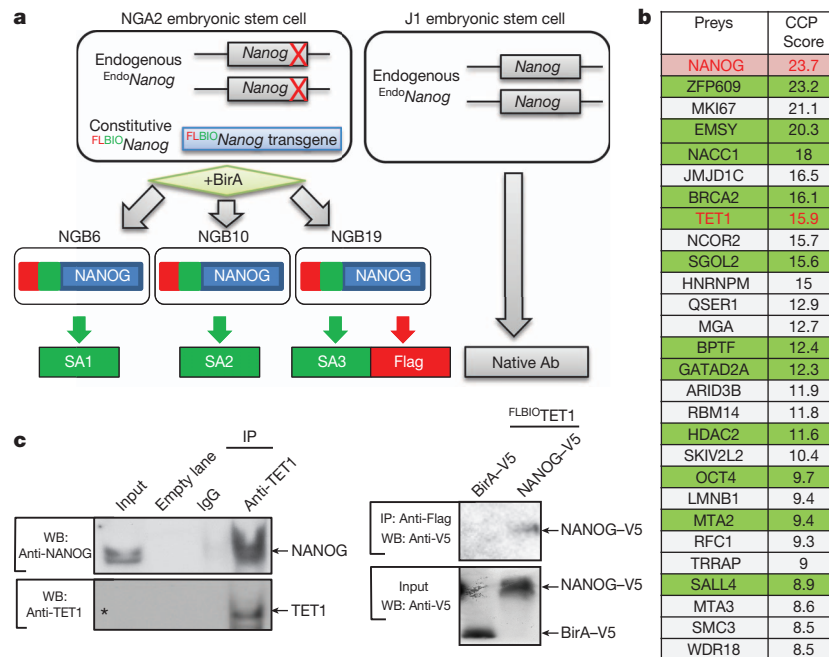
The physical association of TET1 with NANOG prompted us to consider whether TET1 may modulate NANOG function in establishing pluripotency. NANOG and TET1 are only minimally expressed in reprogramming intermediates resulting from retroviral infection of neural stem cells with the reprogramming factors OCT4, KLF4 and c-Myc (rOKM) (Fig. 2a). We addressed whether NANOG-mediated reprogramming of these cells requires functional contribution of TET1. A clonal line of reprogramming intermediates was transfected with a PiggyBac NANOG transgene followed by addition of short interfering RNA (siRNA) against *Tet1* (Supplementary Fig. 9a, b). Downregulation of TET1 reduced NANOG reprogramming efficiency by 26-fold compared with the non-targeting control (Fig. 2b and Supplementary Fig. 9c), indicating that TET1 and/or its associated catalytic activity may be a limiting factor for reprogramming by NANOG.

We then asked if ectopic TET1 expression could enhance NANOG reprogramming activity. Neural stem cells + rOKM were transfected with PiggyBac vectors expressing NANOG, TET1 or TET1 bearing two mutations in the catalytic domain (TET1(H1671Y, D1673A), hereafter TET1Mut)<sup>10</sup> (Supplementary Fig. 10a). Individual expression of wild-type TET1 (TET1WT) or TET1Mut did not have a significant effect on generation of green fluorescent protein (GFP)-positive colonies in cells containing a GFP gene driven by *Oct4* regulatory sequences (Fig. 2c and Supplementary Fig. 10b, c). In contrast, NANOG expression enhanced the generation of iPS cell colonies by more than tenfold (Fig. 2c and Supplementary Fig. 10b, c), in accordance with previous studies<sup>11,14</sup>. Importantly, NANOG-mediated reprogramming efficiency was further augmented by up to fourfold in the

<sup>1</sup>Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK. <sup>2</sup>Department of Developmental and Regenerative Biology, Black Family Stem Cell Institute, Graduate School of Biological Sciences, Mount Sinai School of Medicine, New York, New York 10029, USA. <sup>3</sup>Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK. <sup>4</sup>Epigenetics Programme, The Babraham Institute, Cambridge CB22 3AT, UK. <sup>5</sup>Cambridge Centre for Proteomics, Cambridge Systems Biology Centre, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK. <sup>6</sup>Department of Internal Medicine, and Program in Molecular and Cellular Biology, University of Iowa, Iowa City, Iowa 52242, USA. <sup>7</sup>Department of Pediatrics, Indiana University, Indianapolis, Indiana 46202, USA. <sup>8</sup>Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK.

\*These authors contributed equally to this work.





**Figure 1 | Identification of TET1 as a novel partner of NANOG.**

**a**, Schematic depiction of embryonic stem cells expressing NANOG with Flag (FL) and Biotin (BIO) tags (left), and NANOG antibody (Ab)-based affinity purification (right). SA1–SA3, three independent streptavidin pull-down experiments. **b**, List of 27 preys identified as true interactors ordered by combined cumulative probability (CCP) score. Candidates shaded in green are the ones whose interaction with NANOG has been validated previously<sup>5,25,26</sup> or in this study by immunoprecipitation/co-immunoprecipitation. Two

previously identified NANOG partners, DAX1 and ZFP281<sup>5</sup>, were identified by mass spectrometry (Supplementary Fig. 5a), but not selected as high-confidence interactors using our stringent criteria. **c**, Validation of NANOG–TET1 interaction by immunoprecipitation/co-immunoprecipitation in embryonic stem (left) and HEK293T (right) cells. The asterisk indicates the presence of TET1 in input that can be visualized under longer exposure (Supplementary Fig. 5c). IP, immunoprecipitation; WB, western blot.

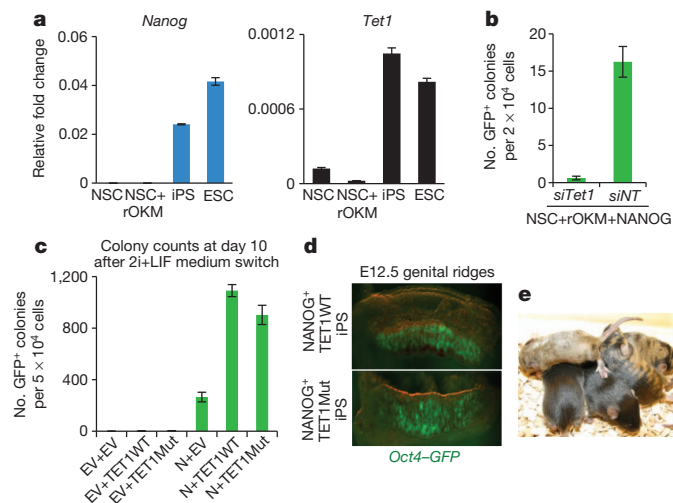
presence of TET1WT transgene, a synergistic effect that is highly reproducible (Fig. 2c and Supplementary Fig. 10b–e). A similar reprogramming synergy was also observed for the combination of NANOG with TET1Mut (Fig. 2c and Supplementary Fig. 10b–e). iPS cells derived with NANOG and either TET1WT or TET1Mut transgenes contributed to the germ lineage and live-born chimaeras following blastocyst injection (Fig. 2d, e). Together, our data show that NANOG and TET1 enhance the efficiency of somatic cell reprogramming in a cooperative manner. This conclusion was corroborated in mouse embryonic fibroblasts, where the combined action of NANOG and TET1WT increased reprogramming efficiency by up to 16-fold (Supplementary Fig. 11).

To explore the molecular mechanism underlying the NANOG–TET1 partnership during reprogramming, we quantified global 5-hydroxymethylcytosine (5hmC) levels<sup>15</sup>. As expected, 5hmC levels were increased upon TET1WT but not TET1Mut expression in neural stem cells + rOKM (Fig. 3a, left). Unexpectedly, co-expression of NANOG and either TET1WT or TET1Mut resulted in increased 5hmC levels (Fig. 3a, left). These results indicate that NANOG can potentiate 5hmC modifications by its association with TET1, and that transcriptional activation of endogenous TET1 and/or its paralogue TET2 may compensate for the lack of catalytic activity of TET1Mut during reprogramming with NANOG. Indeed, TET2 was upregulated by NANOG and TET1WT or TET1Mut, and its expression levels follow a very similar trend to that of 5hmC/C levels (Fig. 3a and Supplementary Fig. 10f, g). TET2 was identified in two out of three independent affinity purifications in our NANOG interactomics study (Supplementary Fig. 12a, b), and physical association of TET2 with NANOG was confirmed by immunoprecipitation/co-immunoprecipitation (Supplementary Fig. 12c, d). TET2 was recently found to contribute to an epigenetic program that directs subsequent transcriptional induction at the pluripotency loci *Nanog* and *Esrrb* during the early stage of somatic cell reprogramming<sup>16</sup>. TET1 and TET2 share the common carboxy-terminal catalytic domain but are divergent in their

amino termini for a CXXC DNA-binding domain, which renders TET2 functionally similar to a truncated form of TET1, TET1C (Fig. 3b). We investigated whether the catalytic activity of TET1 is sufficient to enhance NANOG-mediated reprogramming. Indeed, TET1C acts together with NANOG to enhance reprogramming (Fig. 3c) and retains its physical association with NANOG (Fig. 3d). Not surprisingly, we also observed reprogramming synergy between NANOG and TET2 (Fig. 3c and Supplementary Fig. 13). TET1 and TET2 function is redundant in the context of NANOG-induced reprogramming, as exogenously expressing both TET enzymes together with NANOG does not enhance somatic cell reprogramming beyond expressing NANOG with either individual enzyme (Fig. 3c).

Given that endogenous TET2 was upregulated in the presence of NANOG and TET1Mut (Fig. 3a right and Supplementary Fig. 10f), and that TET2 synergizes with NANOG during reprogramming (Fig. 3c), we tested whether knockdown of *Tet2* could abrogate the reprogramming synergy of NANOG and TET1. Indeed, siRNAs directed against *Tet2* diminished the reprogramming synergy of NANOG and TET1Mut, but did not affect that of NANOG and TET1WT (Fig. 3e, f). This result confirms that TET2 activation compensates for the lack of catalytic activity of TET1Mut during reprogramming with NANOG. Together, our results demonstrate that neither TET1 nor TET2 is sufficient for the induction of pluripotency (Fig. 2c, Supplementary Figs 10 and 11, and Fig. 3c), but either enzyme can partner with NANOG to enhance reprogramming of somatic cells to naive pluripotency.

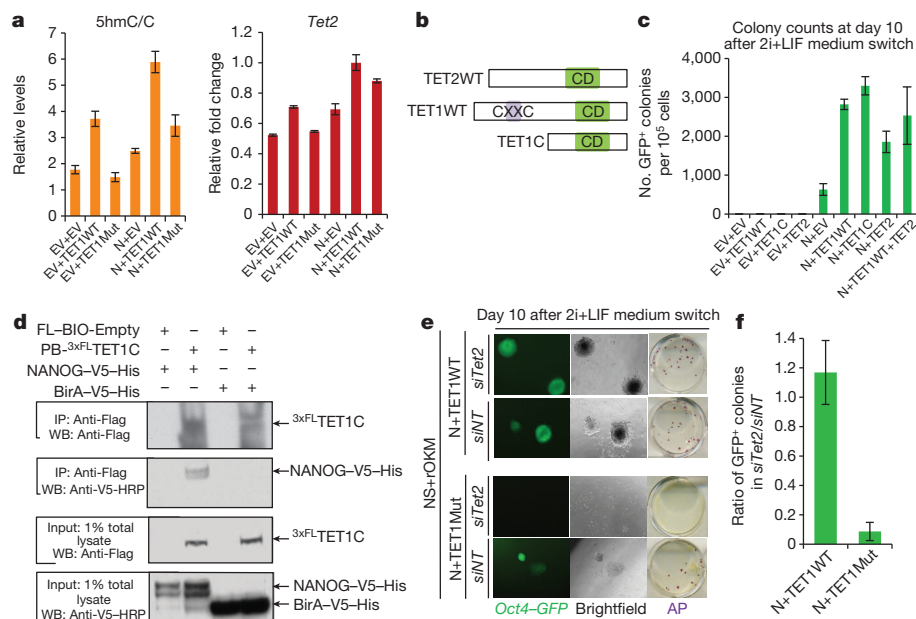
We compared deposited chromatin immunoprecipitation coupled with DNA sequencing (ChIP-Seq) data for both NANOG<sup>17,18</sup> and TET1<sup>19,20</sup> and found a statistically significant overlap between NANOG and TET1 binding sites in the mouse embryonic stem cell genome ( $P < 2 \times 10^{-4}$ , permutation test) (Fig. 4a and Supplementary Table 3). Gene ontology analysis revealed that genes with roles in ‘multicellular organismal development’ and ‘positive regulation of transcription from Pol II promoter’ are enriched in the common



**Figure 2 | Synergy between NANOG and TET1 during reprogramming.** **a**, NANOG and TET1 are specifically expressed in pluripotent cells. ESC, embryonic stem cells; NSC, adult neural stem cells; +rOKM, transduced with retroviral *Oct4*, *Klf4* and *c-Myc* transgenes. **b**, Knockdown of TET1 compromises reprogramming activity of a constitutive NANOG transgene in reprogramming intermediates. *siNT*, non-targeting siRNA control; *siTet1*, siRNA against *Tet1*. Sequences for quantitative PCR primers and siRNA targets are listed in Supplementary Table 5. **c**, Both wild-type and mutant TET1 enhance NANOG-dependent reprogramming. Quantification of the number of iPS colonies at day 10 of 2i+LIF (dual inhibition of mitogen-activated protein kinase signalling (PD0325901, 1  $\mu$ M) and glycogen synthase kinase-3 (GSK3) (CHIR99021, 3  $\mu$ M) with leukaemia inhibitory factor) treatment is shown in Supplementary Fig. 10b. EV, empty vector; N, NANOG. **d**, **e**, Contribution of iPS cells generated with NANOG and TET1WT (top) or TET1Mut (bottom) transgenes to the germline at E12.5 (**d**) and live-born chimaeras (**e**). *Oct4-GFP*, *GFP* gene driven by *Oct4* regulatory sequences. Error bars indicate standard deviation ( $n = 3$ ).

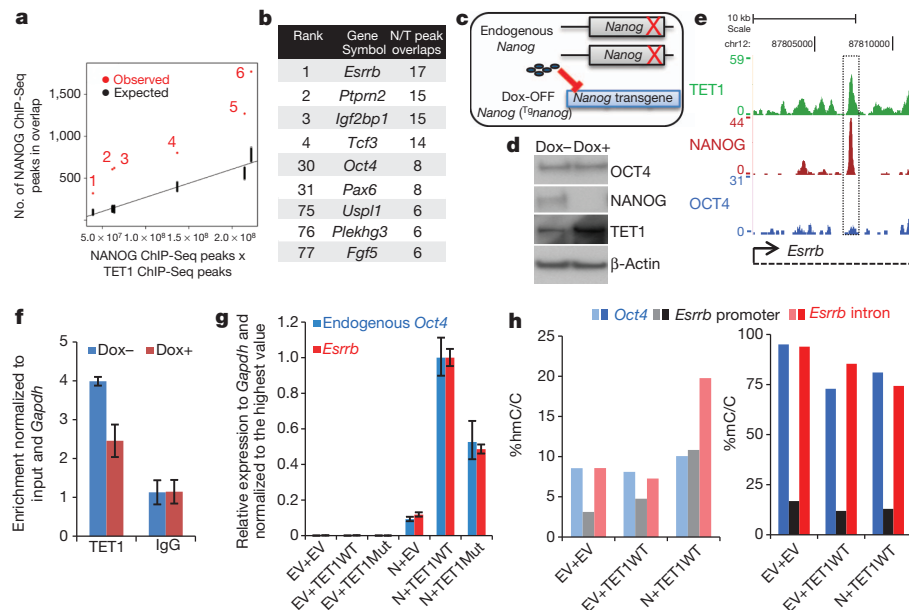
targets (Supplementary Fig. 14). We ranked the common target genes of NANOG and TET1 based on the number of overlapping ChIP-Seq peaks in four studies (Fig. 4b and Supplementary Table 4). Among the common targets with the highest number of overlapping NANOG and TET1 peaks was *Esrrb* (Fig. 4b). To investigate whether NANOG may be required to direct TET1 to shared target genes, we used embryonic stem cells containing an inducible *Nanog* transgene in a *Nanog*<sup>-/-</sup> background<sup>21</sup> (Fig. 4c, d). Loss of NANOG expression reduced TET1 binding to a number of common targets, including *Esrrb* (Fig. 4e, f and Supplementary Fig. 15). NANOG-dependent binding of TET1 to the *Esrrb* locus seems to be independent of OCT4, as OCT4 is not present at the same genomic location (Fig. 4e). Thus, NANOG is responsible for the recruitment of TET1 to a subset of shared genomic loci that are implicated in both the regulation of pluripotency (for example, *Esrrb*) and lineage commitment (for example, *Pax6*). Such NANOG-dependent target binding of TET1 is highlighted by the fact that the truncated form of TET1 lacking the CXXC DNA binding domain (that is, TET1C) maintains its physical interaction and reprogramming synergy with NANOG (Fig. 3b–d).

Because the functional synergy between NANOG and TET enzymes was dependent on catalytic activity, we examined 5hmC levels at NANOG/TET1 peaks in mouse embryonic stem cells. A recent study reported 5hmC enrichment at promoter-distal NANOG binding sites in human embryonic stem cells<sup>22</sup> (Supplementary Fig. 16a). In contrast, we observed an inverse correlation between 5hmC and NANOG/TET1 binding at actively expressed target genes in mouse embryonic stem cells (Supplementary Fig. 16). This led us to consider whether 5hmC may be transiently deposited to common NANOG/TET1 targets before the establishment of pluripotency, that is, during *in vitro* reprogramming when NANOG is required<sup>3</sup>. We focused on target gene regulation of *Esrrb* and *Oct4*, two key pluripotency genes that are among top ranked common targets of NANOG and TET1 (Fig. 4b). Significantly, we observed expression priming of both *Esrrb* and *Oct4* by combined expression of NANOG with TET1WT, TET1Mut, TET1C or TET2 in reprogramming intermediates of two



**Figure 3 | Synergy between NANOG and TET1/TET2 during reprogramming is dependent upon catalytic activity of TET1/TET2.** **a**, Measurement of global levels of 5hmC relative to unmodified C (left, 5hmC/C) and TET2 expression (right) in reprogramming intermediates transfected with PiggyBac (PB) transgenes. **b**, Schematic depiction of wild-type TET1 and TET2 (TET1WT and TET2WT, respectively), and the truncated TET1 mutant (TET1C). Note the absence of a CXXC DNA binding domain in TET2WT and

TET1C proteins. CD, catalytic domain. **c**, Quantification of GFP<sup>+</sup> iPS colonies. **d**, Physical association of NANOG with TET1C. Co-immunoprecipitation was performed in HEK293T cells. HRP, horseradish peroxidase. **e**, TET2 knockdown (*siTet2*) reduces reprogramming efficiency in intermediate cells transgenic for NANOG+TET1Mut compared to NANOG+TET1WT. Non-targeting siRNA (*siNT*) serves as a control. **f**, Quantification of the number of iPS colonies in (**e**). Error bars indicate standard deviation ( $n = 3$ ).



**Figure 4 | Mechanism and genome-wide significance of the NANOG–TET1 interaction.** **a**, Scatterplot showing the observed against expected overlap in genomic binding sites of NANOG and TET1 according to comparisons performed in Supplementary Table 3. **b**, Ranked list of common targets of NANOG and TET1 based on the comparisons in Supplementary Table 3. **c**, Schematic representation of embryonic stem cells harbouring a doxycycline (Dox)-suppressible NANOG transgene in a *Nanog*<sup>-/-</sup> genetic background<sup>21</sup>. **d**, Western blot analysis of OCT4, NANOG, and TET1 expression in NgCGO

embryonic stem cells treated with (+) or without (-) Dox. **e**, Overlapping peaks of TET1 and NANOG from ChIP-Seq studies<sup>17–19</sup> in the *Esrrb* locus. **f**, Relative enrichment of TET1 in the absence (-) and presence (+) of Dox in the *Esrrb* genomic locus as shown in (**e**). **g**, Transcriptional priming of *Esrrb* and *Oct4* by NANOG and TET1 in reprogramming intermediates. **h**, Relative enrichment of 5hmC and 5mC in the *Esrrb* and *Oct4* loci. Error bars indicate standard deviation (n = 3). Primers used for glucosylated hydroxymethyl-sensitive qPCR are listed in Supplementary Table 6.

independent cellular systems (Fig. 4g and Supplementary Figs 11g and 13b). More importantly, we detected increased 5hmC and decreased 5-methylcytosine (5mC) levels at these loci when NANOG is co-expressed with TET1 (Fig. 4h). Thus, NANOG and TET1 act before the transition to naive pluripotency by inducing local transcriptional changes in shared target genes that are critically involved in the regulation of pluripotency.

In summary, we identified 5mC hydroxylases TET1 and TET2 as novel interaction partners of NANOG. TET1/TET2 and NANOG synergistically enhance the efficiency of reprogramming and this phenotype is dependent on the hydroxylation of 5mC to 5hmC during somatic cell reprogramming. This study thus provides mechanistic insight into how NANOG establishes pluripotency, demonstrating that interactions between NANOG and epigenetic regulators fine-tune induced pluripotency. Future experimental work is needed to delineate the precise composition of NANOG–TET1/TET2 protein complexes, and the contribution of other interaction partners to the reprogramming mechanism described herein. Our work supports an emerging view that TET proteins can overcome epigenetic roadblocks during reprogramming and transdifferentiation<sup>16,23</sup>.

## METHODS SUMMARY

**Affinity purification coupled with mass spectrometry (AP-MS).** Nuclear extraction and affinity purification of Flag–biotin–NANOG-associated protein complexes were performed as previously described<sup>5</sup>, with several modifications as described<sup>7</sup>. Three biological replicates were performed for streptavidin agarose-based affinity purification and one each for Flag and NANOG antibody-based affinity purifications. Liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) was used by the Taplin Biological Mass Spectrometry Facility at Harvard Medical School to sequence and identify NANOG affinity purification samples.

**Reprogramming assays.** To investigate the consequences of NANOG and TET1 co-expression during reprogramming, adult neural stem cells were infected with pMX-based retroviral reprogramming factors<sup>24</sup>. Cultures were switched to embryonic stem cell medium (serum/leukaemia inhibitory factor (LIF)) at day 3 post-transduction. A clonal line of proliferative, *Oct4*-GFP negative cells (reprogramming

intermediates) was transfected using nucleofection (Amaxa) with various combinations of NANOG and TET1 PiggyBac transgenes. Selection for stable transgene expression was applied to transfectants for a minimum of 12 days and maintained until medium switch to 2i (dual inhibition of mitogen-activated protein kinase signalling (PD0325901, 1 μM) and glycogen synthase kinase-3 (GSK3) (CHIR99021, 3 μM)) + LIF. Puromycin selection for an *Oct4*-GFP-IRES-puro reporter transgene was applied from day 6 of 2i + LIF treatment. GFP-positive colonies were scored at day 10. Similar reprogramming assays were applied to *Nanog*-GFP-IRES-puro reporter mouse embryonic fibroblasts with modifications described in Supplementary Information.

Received 10 February 2012; accepted 22 January 2013.

Published online 10 February 2013.

- Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643–655 (2003).
- Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642 (2003).
- Silva, J. *et al.* Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722–737 (2009).
- Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
- Wang, J. *et al.* A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364–368 (2006).
- Kim, J., Cantor, A. B., Orkin, S. H. & Wang, J. Use of *in vivo* biotinylation to study protein–protein and protein–DNA interactions in mouse embryonic stem cells. *Nature Protocols* **4**, 506–517 (2009).
- Ding, J., Xu, H., Faiola, F., Ma'ayan, A. & Wang, J. Oct4 links multiple epigenetic pathways to the pluripotency network. *Cell Res.* **22**, 155–167 (2011).
- Rees, J. S. *et al.* *In vivo* analysis of proteomes and interactomes using parallel affinity capture (IPAC) coupled to mass spectrometry. *Mol. Cell Proteomics* <http://dx.doi.org/10.1074/mcp.M110.002386> (2011).
- Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
- Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
- Theunissen, T. W. *et al.* Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. *Curr. Biol.* **21**, 65–71 (2011).
- Dawlaty, M. M. *et al.* Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**, 166–175 (2011).



13. Pereira, C. F. & Fisher, A. G. Heterokaryon-based reprogramming for pluripotency. *Curr. Protoc. Stem Cell Biol.* **9**, 4B.1.1–4B.1.14 (2009).
14. Hanna, J. *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**, 595–601 (2009).
15. Ficiz, G. *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398–402 (2011).
16. Doege, C. A. *et al.* Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* **488**, 652–655 (2012).
17. Chen, X. *et al.* Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106–1117 (2008).
18. Marson, A. *et al.* Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134**, 521–533 (2008).
19. Williams, K. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343–348 (2011).
20. Wu, H. *et al.* Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**, 389–393 (2011).
21. Das, S., Jena, S. & Levasseur, D. N. Alternative splicing produces Nanog protein variants with different capacities for self-renewal and pluripotency in embryonic stem cells. *J. Biol. Chem.* **286**, 42690–42703 (2011).
22. Yu, M. *et al.* Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* **149**, 1368–1380 (2012).
23. Kallin, E. M. *et al.* Tet2 facilitates the derepression of myeloid target genes during CEBP $\alpha$ -induced transdifferentiation of pre-B cells. *Mol. Cell* **48**, 266–276 (2012).
24. Silva, J. *et al.* Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol.* **6**, e253 (2008).
25. Wang, J., Levasseur, D. N. & Orkin, S. H. Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. *Proc. Natl Acad. Sci. USA* **105**, 6326–6331 (2008).
26. Liang, J. *et al.* Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nature Cell Biol.* **10**, 731–739 (2008).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank W. Mansfield for blastocyst injections, A. Radzishewska for cell culture assistance, and R. Jaenisch for *Tet1*<sup>-/-</sup> embryonic stem cells. This study was supported by a grant from the NIH (1R01-GM095942-01A1), a grant from New York state Department of Health (NYSTEM#C026420), and a seed fund from the Black Family Stem Cell Institute to J.W., by the Wellcome Trust Fellowship (WT086692MA) and the Isaac Newton Trust Grant (11.19(ad)) to J.C.R.S., who is a Wellcome Trust Career Development Fellow, by the BBSRC, the MRC, the Wellcome Trust, and EU Epigenesys and BLUEPRINT to W.R., and by the Wellcome Trust Fellowship WT079249 to T.W.T.

**Author Contributions** J.C.R.S. and J.W. conceived the project, designed the experiments, prepared and approved the manuscript. T.W.T. designed and performed experiments and wrote the manuscript draft. Y.C., J.D., F.F., M.F. and A.S. designed and performed experiments and prepared the manuscript. T.A.H. designed and performed experiments. P.V.S. performed interactomics data analysis. M.L. provided technical assistance. S.D. provided bioinformatic analysis. S.Da., D.N.L., Z.L. and M.X. contributed to the reagents. W.R. designed experiments and contributed to the reagents. J.C.R.S. and J.W. are equal senior and corresponding authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.W. ([jianlong.wang@mssm.edu](mailto:jianlong.wang@mssm.edu)) or J.C.R.S. ([jcs64@cam.ac.uk](mailto:jcs64@cam.ac.uk)).

# Multiple phases of chondrocyte enlargement underlie differences in skeletal proportions

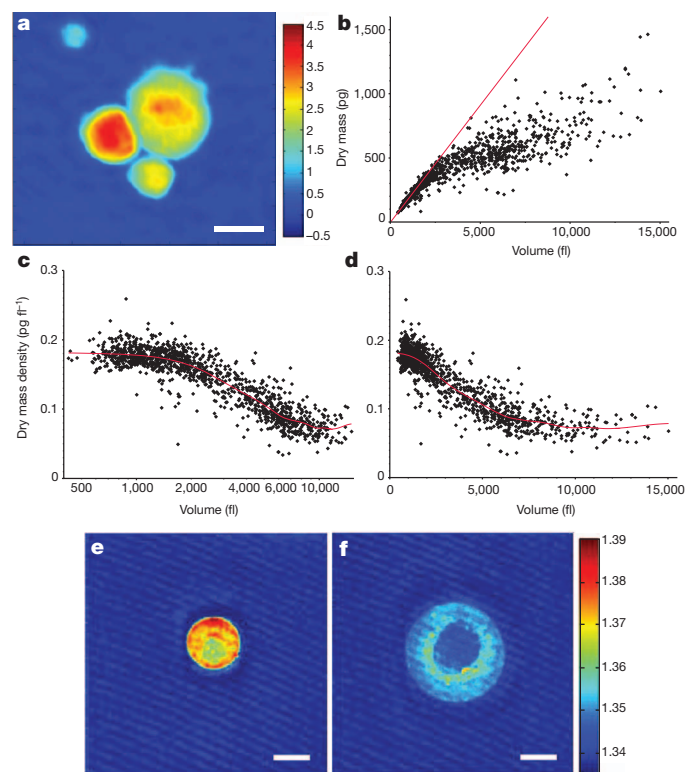
Kimberly L. Cooper<sup>1\*</sup>, Seungeun Oh<sup>2\*</sup>, Yongjin Sung<sup>3</sup>, Ramachandra R. Dasari<sup>3</sup>, Marc W. Kirschner<sup>2</sup> & Clifford J. Tabin<sup>1</sup>

The wide diversity of skeletal proportions in mammals is evident upon a survey of any natural history museum's collections and allows us to distinguish between species even when reduced to their calcified components. Similarly, each individual is comprised of a variety of bones of differing lengths. The largest contribution to the lengthening of a skeletal element, and to the differential elongation of elements, comes from a dramatic increase in the volume of hypertrophic chondrocytes in the growth plate as they undergo terminal differentiation<sup>1–7</sup>. However, the mechanisms of chondrocyte volume enlargement have remained a mystery<sup>8–11</sup>. Here we use quantitative phase microscopy<sup>12</sup> to show that mammalian chondrocytes undergo three distinct phases of volume increase, including a phase of massive cell swelling in which the cellular dry mass is significantly diluted. In light of the tight fluid regulatory mechanisms known to control volume in many cell types<sup>13</sup>, this is a remarkable mechanism for increasing cell size and regulating growth rate. It is, however, the duration of the final phase of volume enlargement by proportional dry mass increase at low density that varies most between rapidly and slowly elongating growth plates. Moreover, we find that this third phase is locally regulated through a mechanism dependent on insulin-like growth factor. This study provides a framework for understanding how skeletal size is regulated and for exploring how cells sense, modify and establish a volume set point.

Each of the long bones initially forms in the embryo as a similarly sized cartilage rudiment that only subsequently undergoes differential regulation of growth. The elongation of a skeletal element occurs at the growth plate<sup>14,15</sup>, each consisting of three distinct zones: resting round chondrocytes near the end of an element give rise to clonal columns of flattened proliferating chondrocytes that then terminally differentiate into hypertrophic chondrocytes nearest the bony centre of an element. Although multiple cellular parameters contribute to lengthening of skeletal elements—including proliferation, matrix deposition and hypertrophic cell enlargement—the greatest contribution to growth rate in mammals is due to the massive volume enlargement of hypertrophic chondrocytes expanding the skeletal tissue in the direction of longitudinal growth within laterally restricting matrix channels<sup>1–3</sup>. In addition to being the largest contributor to the elongation rate of a given skeletal element, this parameter is largely responsible for the difference in growth rates between different skeletal elements within an individual, between homologous elements in different species, and within a single element as an animal ages<sup>4–7</sup>. Surprisingly, however, given its critical importance in determining the growth rate of each bone and the overall stature of the individual, the mechanism is poorly understood. It even remains unclear whether hypertrophic chondrocyte volume increases by true hypertrophy, maintaining constant density during growth through an increase in macromolecules and organelles, or by cell swelling through disproportionate fluid uptake, which is ordinarily a hallmark of disease<sup>13</sup>.

To determine whether hypertrophic chondrocytes enlarge by cellular hypertrophy and/or swelling, we applied methods of diffraction

phase microscopy to measure the dry mass of individual unstained live cells dissociated from growth plate cartilage. (Fig. 1a and Supplementary Methods)<sup>16</sup>. Together with volume information about the sample, calculated here based on a well-supported spherical approximation for dissociated chondrocytes (Supplementary Methods), this allows the calculation of dry mass density. Diffraction phase microscopy measurements of a variety of cell types consistently reflect a 'normal' dry mass density for healthy living cells at approximately 0.182 pg per femtolitre (fl) in agreement with the concentration of cytoplasm previously determined by index matching in human oral epithelial cells<sup>17</sup>.



**Figure 1 | Hypertrophic chondrocytes increase in volume through three distinct phases including a phase of massive cell swelling.** **a**, Quantitative phase image of dissociated mouse proximal tibia hypertrophic chondrocytes at postnatal day 5 (P5). Colour bar represents phase shift in radians. **b**, Volume versus dry mass plotted for individual chondrocytes. Linear regression for cells up to 1,000 fl highlights divergence of larger cells from an initial slope of 0.183 pg fl<sup>-1</sup>. **c**, **d**, Log scale (**c**) and linear scale (**d**) plots of volume versus dry mass density. Lambda for the smoothed spline in (**c**) and (**d**) is  $5 \times 10^9$ .  $R^2$  value is 0.84.  $n = 1,249$  cells summed across five independent experiments. **e**, **f**, Horizontal cross sections from regularized tomographic phase microscopy density reconstructions of a small (**e**) and large (**f**) mouse tibia chondrocyte. Colour bar represents refractive index and thus dry mass density. Scale bars, 10  $\mu$ m.

<sup>1</sup>Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>2</sup>Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>3</sup>George R. Harrison Spectroscopy Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

\*These authors contributed equally to this work.

This includes maturing megakaryocytes that reach volumes comparable to the largest hypertrophic chondrocytes and ten times the average somatic cell volume (Supplementary Fig. 1).

In contrast, analysis of chondrocytes from the rapidly elongating mouse proximal tibia reveals that there are three distinct phases of hypertrophic cell enlargement. In phase 1, an initial increase of about threefold from approximately 600 fl to 2,000 fl is characterized by true hypertrophy—a proportionate increase in dry mass production and fluid uptake thus maintaining the normal dry mass density at  $0.183 \text{ pg fl}^{-1}$  (Fig. 1b, c). In phase 2, a fourfold enlargement from about 2,000 fl to 8,000 fl is characterized by cell swelling. Volume increases at a rate disproportionate to the continuing rate of dry mass production resulting in a dramatic dilution of dry mass density to approximately  $0.07 \text{ pg fl}^{-1}$  (Fig. 1c). In phase 3, at volumes larger than 8,000 fl, the dry mass density once again stabilizes, and cells continue to enlarge another twofold to about 14,000 fl by proportionately increasing dry mass and fluid volume at this lower density (Fig. 1d). Swelling in phase 2 allows cells to reach volumes two to three times greater in phase 3 than if they relied entirely on the proportionate increase in dry mass at high density (Fig. 1b, linear regression).

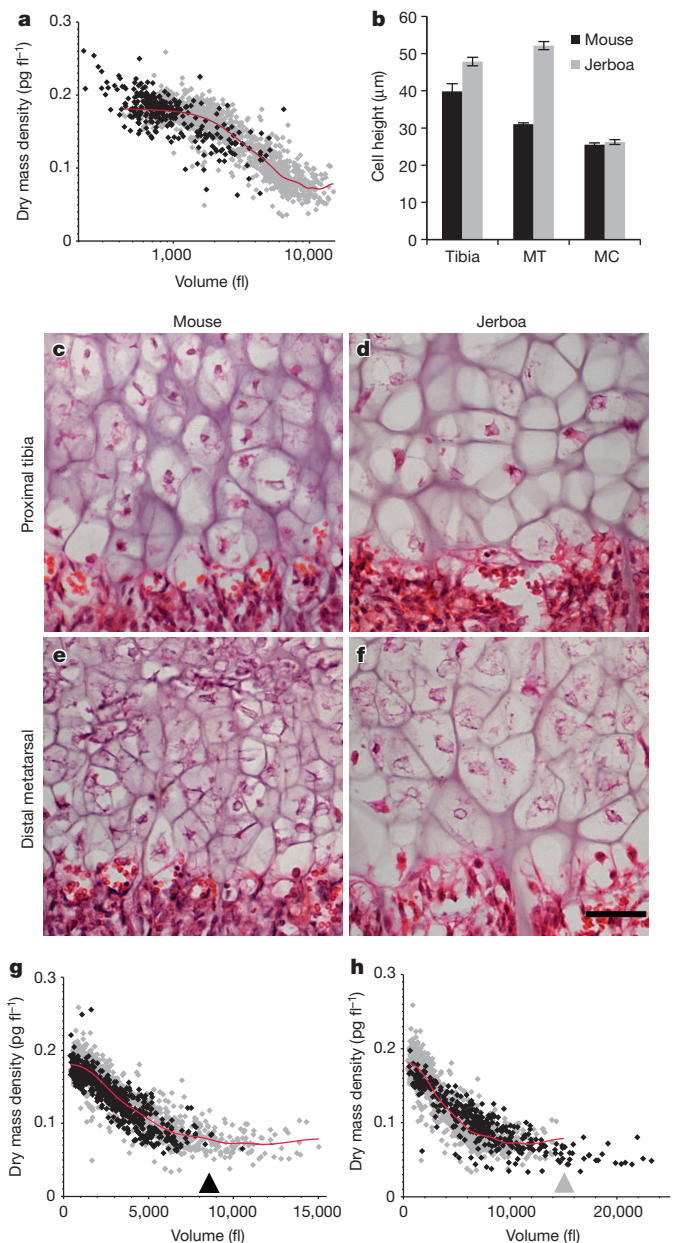
To verify that this decrease in density is characteristic of hypertrophic differentiation, we imaged a subset of the small high-density cells and large low-density cells using regularized tomographic phase microscopy to generate a refractive index map of dry mass density in three-dimensions (Supplementary Methods). This independent approach confirms that the largest cells reduce their dry mass density by approximately 60%, and moreover indicates that dry mass is low throughout the cytoplasm with a slightly higher density ring around the nucleus (Fig. 1e, f and Supplementary Fig. 14).

Understanding the cellular process by which hypertrophic cells enlarge provides a framework for considering how that process is modulated to achieve differential growth of individual elements within a species and of homologous elements between species. In contrast to the large hypertrophic chondrocytes of the rapidly elongating proximal tibia, the slowly elongating proximal radius has much smaller hypertrophic chondrocytes. We find that these cells go through phase 1 and enter phase 2, similar to the cells of the proximal tibia, reaching a volume of approximately 5,000 fl. However, they truncate the remainder of phase 2 after dilution of dry mass density to approximately  $0.10 \text{ pg fl}^{-1}$  and completely eliminate phase 3 (Fig. 2a).

We next compared growth plates of the mouse to those of the lesser Egyptian jerboa, *Jaculus jaculus*, a small bipedal rodent with greatly elongated hindlimbs. In particular, the metatarsals of jerboa feet rapidly elongate during early postnatal development to approximately 2.5 times the relative proportion of mouse metatarsals<sup>18</sup>. The mouse distal metatarsal growth plate is intermediate in its growth rate and hypertrophic chondrocyte size compared to the mouse proximal tibia and radius (Fig. 2b and Supplementary Fig. 2). Although other aspects of the jerboa distal metatarsal growth plate are also altered, including the total cell number in each zone (Supplementary Fig. 3), the height of individual hypertrophic chondrocytes is increased by 58% compared to the metatarsal of the mouse, suggesting a significant contribution of hypertrophic chondrocyte volume to the increased rate of growth of this element (Fig. 2b, e, f). In contrast, the jerboa proximal tibia hypertrophic chondrocytes are only slightly larger than their counterparts in the mouse (Fig. 2b–d). Unlike the metatarsals, the metacarpals of the jerboa forelimb, as well as other bones of the forelimb, are similar in size and proportions to those of the mouse with hypertrophic chondrocytes of comparable size (Fig. 2b).

The hypertrophic chondrocytes of the jerboa proximal tibia show very similar growth properties to those of the mouse when examined by diffraction phase microscopy, including all three phases of volume enlargement (Supplementary Fig. 4). However, there is a striking difference between the metatarsal chondrocytes of the two species. Although mouse metatarsal hypertrophic chondrocytes are indeed intermediate in size between the proximal radius and proximal tibia,

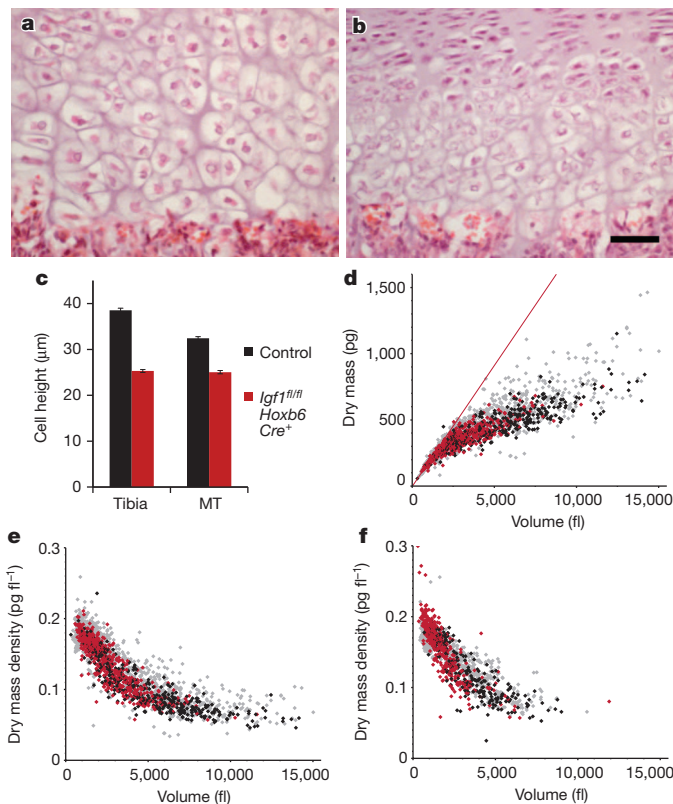
reaching a maximum volume of about 8,000 fl by completing phases 1 and 2 and truncating phase 3 (Fig. 2g), hypertrophic chondrocytes of the jerboa metatarsals increase almost 40-fold from their initial volume to approximately 23,000 fl, greater than the volume of tibia chondrocytes in either species (Fig. 2h). This is accomplished by following the



**Figure 2 | Differences in cell size associated with different skeletal growth rates are attributed to modulating a common growth trajectory.**

**a**, Dissociated postnatal day 5 (P5) mouse proximal radius chondrocytes (black,  $n = 292$  cells) compared to proximal tibia chondrocytes (grey in all panels, data from Fig. 1). The  $x$  axis is in log scale. **b**, Quantification of average maximum cell height  $\pm$  s.e.m. for  $n = 3$  animals of each species at P7 ( $>50$  cells per growth plate). Two-tailed student's  $t$ -test shows significant differences in the tibia and metatarsal growth plates between the two species ( $P < 10^{-5}$ ) but not in the metacarpals ( $P = 0.978$ ). MC, metacarpal growth plate; MT, metatarsal growth plate. **c–f**, Histological comparison of mouse and jerboa proximal tibia and distal metatarsal hypertrophic zones at P7. Scale bar, 50  $\mu\text{m}$ . **g**, Dissociated P5 mouse metatarsal chondrocytes (black,  $n = 634$  cells) compared to mouse tibia chondrocytes. Black arrowhead approximates the end of the metatarsal distribution. **h**, Jerboa metatarsal chondrocytes (black,  $n = 366$  cells) compared to mouse tibia chondrocytes. Grey arrowhead approximates the end of the mouse tibia distribution.

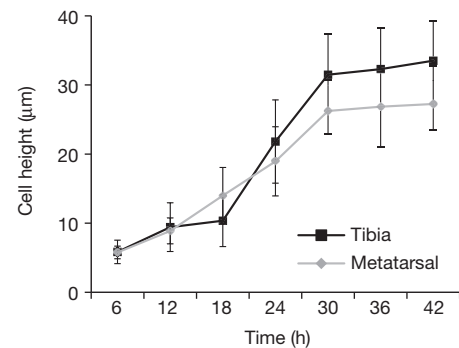




**Figure 3 | *Igf1* is required for phase 3 of volume enlargement by dry mass production at low dry mass density.** **a, b,** Histology of postnatal day 7 (P7) mouse proximal tibia hypertrophic zone of control (**a**, *Igf1*<sup>fl/+</sup>; *Hoxb6-Cre*<sup>+</sup>) and *Igf1* conditional mutant animals (**b**, *Igf1*<sup>fl/fl</sup>; *Hoxb6-Cre*<sup>+</sup>). Scale bar, 50  $\mu$ m. **c,** Bar plot demonstrating an average maximum cell height reduction of 34% in *Igf1* conditional mutant tibia and 23% in metatarsal chondrocytes compared to control littermates. Error bars indicate  $\pm$  s.e.m. for  $n = 3$  animals of each genotype (total = >80 cells per growth plate). **d,** Dry mass versus volume plot of dissociated P5 *Igf1* mutant proximal tibia hypertrophic chondrocytes. Red data points are *Igf1* mutant chondrocytes ( $n = 569$  cells), black are littermate control chondrocytes ( $n = 373$  cells) and grey represent the total wild-type data set for tibia (**d, e**) and metatarsal (**f**) (wild-type data from Figs 1 and 2). **e, f,** Dry mass density versus volume plots for *Igf1* mutant tibia chondrocytes ( $n = 569$  cells) (**e**) and *Igf1* mutant metatarsal chondrocytes ( $n = 412$  cells) (**f**).

same triphasic growth trajectory common to chondrocytes of other growth plates and then extending phase 3 to reach a maximal volume by continued proportionate increase in dry mass and fluid volume at low dry mass density.

Little is known about the molecular mechanisms that control chondrocyte enlargement, or the regulation of final cell size, as few of the mutants affecting skeletal development have been examined for size of individual hypertrophic chondrocytes. An intriguing exception is the null mutation of the gene encoding insulin-like growth factor 1 (*Igf1*)<sup>19</sup>. *Igf1* functions in a variety of target tissues to promote protein synthesis and cell growth<sup>19,20</sup> and is strongly expressed in both proliferating and pre-hypertrophic chondrocytes. *Igf1*-deficient mice are 35% smaller than controls but have the same number of hypertrophic chondrocytes, although each cell is 30% shorter in the direction of elongation, a finding we confirmed in mice where the floxed *Igf1*<sup>tm1Dlr</sup> allele<sup>21</sup> was conditionally deleted from the hindlimb using *Hoxb6-Cre* (ref. 22) (Fig. 3a–c). It is additionally intriguing to note that there is no distinction between the heights of chondrocytes in the proximal tibia and distal metatarsal of this mutant (Fig. 3c), indicating that *Igf1* may play an important role in the establishment of growth-plate-dependent cell size. Although this effect could be modulated by any member of the *Igf1* signalling pathway, evidence suggests a possible role for the receptor, *Igf1R*, which maintains higher levels of



**Figure 4 | Mouse proximal tibia and distal metatarsal hypertrophic chondrocytes rapidly increase in average cell height.** Time course of the average and standard deviation of BrdU labelled cell height indicating the rate of chondrocyte size increase after the last mitotic cycle.  $n > 25$  cells from three individuals for each time point and growth plate.

expression over time in growth plates that continue to elongate at faster rates in maturing mice<sup>23</sup>.

We employed diffraction phase microscopy to determine which phase(s) of volume enlargement are affected by *Igf1*. *Igf1*-deficient hypertrophic chondrocytes undergo normal phase 1 and phase 2 of enlargement, reaching approximately 7,000 fl largely by cell swelling, but do not progress to phase 3 and thus fail to further double their volume by the continued production of dry mass at low density (Fig. 3d–f). Taken together, our results indicate that there are three distinct phases of chondrocyte hypertrophy, and it is regulation of the *Igf1*-dependent third phase that is responsible for much of the variation in skeletal elongation rate.

Finally, previous studies of the neonatal bat and mouse forelimb indicated that the entire hypertrophic zone of each growth plate turns over once in about 24 h regardless of the maximum volume attained by individual chondrocytes, the number of hypertrophic chondrocytes, or the rate of growth plate elongation<sup>7</sup>. This suggests that growth plates elongating at different rates adjust the rate of cell volume increase to fall within a 24-h lifespan constraint. To assess the pace of cellular maturation and enlargement, we marked proliferating mouse chondrocytes with 5-bromodeoxyuridine (BrdU) and followed the progression of the first labelled cells to emerge into the post-mitotic hypertrophic columns onward to the chondro-osseous junction (Supplementary Fig. 5). We find that the rate of cell height increase in the proximal tibia and distal metatarsal growth plates is extremely rapid, more than tripling the height of proximal tibia chondrocytes within approximately 12 h (Fig. 4). Once cells reach their average final height, with a steeper slope in the larger cells of the tibia, they remain at this terminal size in the hypertrophic growth columns for an additional 12 h before turnover at the chondro-osseous junction.

Through coordination of this multiphase process, cell swelling allows chondrocytes to enlarge extraordinarily rapidly while presumably lowering the energetic cost of growth, and volumes are subsequently amplified in the most rapidly elongating skeletal elements by the continued *Igf1*-dependent production of mass. This unique mechanism of volume enlargement suggests that chondrocyte hypertrophy will serve as a valuable model for cell volume homeostasis, in addition to our findings that provide insight into skeletal morphogenesis and evolution.

## METHODS SUMMARY

The distal growth plate adjacent to the chondro-osseous junction was micro-dissected from P5 mouse and jerboa pups and dissociated for approximately 4 h in 2 mg ml<sup>-1</sup> collagenase D at 37 °C. Cells were transferred to a 14-mm diameter glass-bottom dish for diffraction phase imaging. The specifications of the diffraction phase microscope are provided in the Supplementary Methods. The interference image, captured by charge coupled device (CCD), was used to compute a quantitative phase image in MATLAB. See Supplementary Methods for the

computation algorithm. The quantitative phase is converted to dry mass using the specific refractive increment value of  $0.18 \text{ ml g}^{-1}$ . For a subpopulation of cells, sphericity was determined using three-dimensional confocal morphometry and the refractive index contrast method (Supplementary Figs 12 and 13). As dissociated chondrocytes are spherical, we use the volume calculated from measured cell diameter for cells analysed by diffraction phase microscopy. Dry mass density was calculated by dividing the total cell dry mass by its volume. Dry mass density of a subset of cells was confirmed using regularized tomographic phase microscopy (Supplementary Fig. 14). See Supplementary Methods for details.

**Full Methods** and any associated references are available in the online version of the paper.

**Received 18 March 2012; accepted 29 January 2013.**

**Published online 13 March 2013.**

- Wilsman, N. J., Farnum, C. E., Leiferman, E. M., Fry, M. & Barreto, C. Differential growth by growth plates as a function of multiple parameters of chondrocytic kinetics. *J. Orthop. Res.* **14**, 927–936 (1996).
- Hunziker, E. B., Schenk, R. K. & Cruz-Orive, L. M. Quantitation of chondrocyte performance in growth-plate cartilage during longitudinal bone growth. *J. Bone Joint Surg. Am.* **69**, 162–173 (1987).
- Hunziker, E. B. & Schenk, R. K. Physiological mechanisms adopted by chondrocytes in regulating longitudinal bone growth in rats. *J. Physiol. (Lond.)* **414**, 55–71 (1989).
- Breur, G. J., VanEnkevort, B. A., Farnum, C. E. & Wilsman, N. J. Linear relationship between the volume of hypertrophic chondrocytes and the rate of longitudinal bone growth in growth plates. *J. Orthop. Res.* **9**, 348–359 (1991).
- Kuhn, J. L., Delacey, J. H. & Leenellett, E. E. Relationship between bone growth rate and hypertrophic chondrocyte volume in New Zealand white rabbits of varying ages. *J. Orthop. Res.* **14**, 706–711 (1996).
- Wilsman, N. J., Bernardini, E. S., Leiferman, E., Noonan, K. & Farnum, C. E. Age and pattern of the onset of differential growth among growth plates in rats. *J. Orthop. Res.* **26**, 1457–1465 (2008).
- Farnum, C. E., Tinsley, M. & Hermanson, J. W. Forelimb versus hindlimb skeletal development in the big brown bat, *Eptesicus fuscus*: functional divergence is reflected in chondrocytic performance in autopodial growth plates. *Cells Tissues Organs* **187**, 35–47 (2008).
- Buckwalter, J. A., Mower, D., Ungar, R., Schaeffer, J. & Ginsberg, B. Morphometric analysis of chondrocyte hypertrophy. *J. Bone Joint Surg. Am.* **68**, 243–255 (1986).
- Farnum, C. E., Lee, R., O'Hara, K. & Urban, J. P. G. Volume increase in growth plate chondrocytes during hypertrophy: the contribution of organic osmolytes. *Bone* **30**, 574–581 (2002).
- Bush, P. G., Parisinos, C. A. & Hall, A. C. The osmotic sensitivity of rat growth plate chondrocytes in situ: clarifying the mechanisms of hypertrophy. *J. Cell. Physiol.* **214**, 621–629 (2008).
- Bush, P. G., Pritchard, M., Logman, M. Y., Damron, T. A. & Hall, A. C. A key role for membrane transporter NKCC1 in mediating chondrocyte volume increase in the mammalian growth plate. *J. Bone Miner. Res.* **25**, 1594–1603 (2010).
- Barer, R. Interference microscopy and mass determination. *Nature* **169**, 366–367 (1952).
- Hoffmann, E. K., Lambert, I. H. & Pedersen, S. F. Physiology of cell volume regulation in vertebrates. *Physiol. Rev.* **89**, 193–277 (2009).
- Hunziker, E. B. Mechanism of longitudinal bone growth and its regulation by growth plate chondrocytes. *Microsc. Res. Tech.* **28**, 505–519 (1994).
- Kronenberg, H. M. Developmental regulation of the growth plate. *Nature* **423**, 332–336 (2003).
- Popescu, G., Ikeda, T., Dasari, R. R. & Feld, M. S. Diffraction phase microscopy for quantifying cell structure and dynamics. *Opt. Lett.* **31**, 775–777 (2006).
- Barer, R. Determination of dry mass, thickness, solid and water concentration in Living Cells. *Nature* **172**, 1097–1098 (1953).
- Cooper, K. L. The lesser Egyptian jerboa, *Jaculus jaculus*: a unique rodent model for evolution and development. *Cold Spring Harb. Protocols* **2011**, pdb.emo066704 (2011).
- Wang, J., Zhou, J. & Bondy, C. A. Igf1 promotes longitudinal bone growth by insulin-like actions augmenting chondrocyte hypertrophy. *FASEB J.* **13**, 1985–1990 (1999).
- Oldham, S. & Hafen, E. Insulin/IGF and target of rapamycin signaling: a TOR de force in growth control. *Trends Cell Biol.* **13**, 79–85 (2003).
- Yakar, S. Normal growth and development in the absence of hepatic insulin-like growth factor I. *Proc. Natl Acad. Sci. USA* **96**, 7324–7329 (1999).
- Lowe, L. A., Yamada, S. & Kuehn, M. R. HoxB6-Cre transgenic mice express Cre recombinase in extra-embryonic mesoderm, in lateral plate and limb mesoderm and at the midbrain/hindbrain junction. *Genesis* **26**, 118–120 (2000).
- Serrat, M. A., Lovejoy, C. O. & King, D. Age- and site-specific decline in insulin-like growth factor-I receptor expression is correlated with differential growth plate activity in the mouse hindlimb. *Anat. Rec. (Hoboken)* **290**, 375–381 (2007).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We would like to thank T. J. Mitchison, C. E. Farnum and members of the Developmental Bone Morphogenesis program project grant (National Institutes of Health (NIH)) for helpful discussions. We also thank the Nikon Imaging Center at Harvard Medical School for technical support, A. Luyten and R. Shivdasani for providing mouse megakaryocytes and P. Ramirez for jerboa care. This work was supported by NIH grants P01DK056246 to C.J.T.; R01GM026875 to M.W.K.; and by NIH grant P41RR02594, National Science Foundation (NSF) grant DBI0754339 and support from the Hamamatsu Corporation to R.R.D.

**Author Contributions** K.L.C. and S.O. conceived the project and carried out most of the experiments. Y.S. and R.R.D. carried out critical tomographic experiments validating the primary approaches taken. C.J.T. and M.W.K. supervised the project. K.L.C., S.O., M.W.K. and C.J.T. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.L.C. ([kcooper@genetics.med.harvard.edu](mailto:kcooper@genetics.med.harvard.edu)).

## METHODS

**Animals.** CD-1 was chosen as the wild-type mouse strain for this study. *Igf1<sup>tm1Dlr</sup>* conditional mice<sup>21</sup> and *Hoxb6-Cre* transgenic mice<sup>22</sup> were previously described. Jerboas were housed and reared as previously described<sup>24</sup>. All animal protocols were approved by the Harvard Medical Area Standing Committee on Animals.

**Sectioning and histology.** Dissected skeletal elements were fixed overnight at 4 °C in 4% paraformaldehyde and then carried through a graded series of ethanol dehydration washes before transition through xylenes and into paraffin wax. Sections were cut at 10–12-µm thickness and stained with haematoxylin and eosin. Average maximum cell heights were measured in the axis of linear growth through the lacunae surrounding the largest cells with a clear nuclear profile from digital images of the hypertrophic zone and averaged across at least 4 sections from at least 3 individuals. BrdU (100 mg kg<sup>-1</sup>) or oxytetracycline hydrochloride (20 mg kg<sup>-1</sup>) was injected into the peritoneum of P5 mice before harvest. BrdU was detected using a rat anti-BrdU antibody (AbD, Serotech) followed by goat anti-rat Alexa594 (Invitrogen) in paraffin sections. Oxytetracycline was detected by fluorescence in bisected skeletal elements.

**Chondrocyte and megakaryocyte isolation.** P5 animals were chosen for this study because the tibia and metatarsal growth plates are rapidly elongating in both species, but the metatarsal epiphysis (secondary ossification centre) has not yet formed. While the epiphysis of the metatarsal forms by P7 in the mouse, it appears later in the jerboa (Supplementary Fig. 3). As the hypertrophic chondrocytes reside in a small domain nearest the chondro-osseous junction, we enriched for these cells by using a razor blade to remove a majority of the cartilage containing resting and proliferative chondrocytes leaving the cells closest to the chondro-osseous junction and a small amount of the adjacent trabecular bone. Growth plates were bisected longitudinally and incubated for 45 min at 37 °C in 2 mg ml<sup>-1</sup> collagenase D (Roche) in DMEM/F12 (Invitrogen, 290–330 mOsm) plus 10% fetal calf serum. After the initial incubation, the bone collar, trabecular bone and loosened connective tissues were manually removed with forceps, and the remaining cartilage fragments were transferred to a fresh dish of collagenase digestion media. Cells were incubated for an additional 2–3 h with occasional swirling until cells mostly dissociated from the surrounding matrix. Dissociated chondrocytes were transferred to 35-mm dishes with a 14-mm diameter, 1.5-thickness glass bottom (MatTek) and imaged immediately (see Supplementary Methods for details of imaging). The largest hypertrophic chondrocytes are a small population of the most mature cells, therefore we further enriched for this population in our data analysis by scanning for fields of view containing the largest cells in the dish and quantified all of the intact spherical neighbours.

There has been a longstanding discussion in the literature regarding the extracellular osmolarity of chondrocytes, primarily in the articular cartilage<sup>9,10,25,26</sup>. Much of this data are based on the theoretical ionic environment in association with charged glycosaminoglycans according to the Donnan-Gibbs equilibrium, although to our knowledge the extracellular osmolarity of growth plate cartilage has not been directly measured. While serum osmolarity is approximately 280 mOsm, the osmolarity in association with cartilage may be upwards of 400 mOsm. To address the possibility that the swelling we observe in chondrocytes at larger volumes may be the response to media of low osmolarity, we repeated the diffraction phase microscopy measurements in mouse proximal tibia growth plates dissociated in 424 mOsm DMEM/F12 raised with sucrose. Media osmolarity was measured using the Vapro Model 5600 (Wescor). We find the same three phases, including the phase of cell swelling, indicating this is an inherent property of growth plate chondrocytes and not the passive response to an abnormal osmotic environment (Supplementary Fig. 6). We presume that the 4 h from dissection to imaging is enough time for volume regulatory mechanisms to compensate for any response to osmotic stress that may have occurred. Indeed, 4D confocal imaging of *in situ* porcine articular chondrocytes after osmotic stress shows a mean recovery rate of 4.1% ± 1.8% per min with 96% volume recovery after about 12 min<sup>27</sup>.

Fetal megakaryocytes were isolated from embryonic day 14.5 mouse livers according to previously published protocols<sup>28,29</sup> or generously donated by A. Luyten and R. Shivdasani. Cells were imaged by diffraction phase microscopy one day after harvest or at maturity after five days in culture.

24. Jordan, B., Vercammen, P. & Cooper, K. L. Husbandry and breeding of the lesser Egyptian jerboa, *Jaculus jaculus*. *Cold Spring Harb Protocols* **2011**, <http://dx.doi.org/10.1101/pdb.prot066712> (2011).
25. Maroudas, A. & Evans, H. A study of ionic equilibria in cartilage. *Connect. Tissue Res.* **1**, 69–77 (1972).
26. Urban, J. P. G., Hall, A. C. & Gehl, K. A. Regulation of matrix synthesis rates by the ionic and osmotic environment of articular chondrocytes. *J. Cell. Physiol.* **154**, 262–270 (1993).
27. Errington, R. J., Fricker, M. D., Wood, J. L., Hall, A. C. & White, N. S. Four-dimensional imaging of living chondrocytes in cartilage using confocal microscopy: a pragmatic approach. *Am. J. Physiol.* **272**, C1040–C1051 (1997).
28. Lecine, P., Blank, V. & Shivdasani, R. Characterization of the hematopoietic transcription factor NF-E2 in primary murine megakaryocytes. *J. Biol. Chem.* **273**, 7572–7578 (1998).
29. Shivdasani, R. A. & Schulze, H. Culture, expansion, and differentiation of murine megakaryocytes. *Current Protocols Immunol.* <http://dx.doi.org/10.1002/0471142735.im22f06s67> (2005).



# Brown-fat paucity due to impaired BMP signalling induces compensatory browning of white fat

Tim J. Schulz<sup>1†</sup>, Ping Huang<sup>2</sup>, Tian Lian Huang<sup>1</sup>, Ruidan Xue<sup>1,3</sup>, Lindsay E. McDougall<sup>1</sup>, Kristy L. Townsend<sup>1</sup>, Aaron M. Cypess<sup>1</sup>, Yuji Mishina<sup>4</sup>, Emanuela Gussoni<sup>2</sup> & Yu-Hua Tseng<sup>1,5</sup>

Maintenance of body temperature is essential for the survival of homeotherms. Brown adipose tissue (BAT) is a specialized fat tissue that is dedicated to thermoregulation<sup>1</sup>. Owing to its remarkable capacity to dissipate stored energy and its demonstrated presence in adult humans<sup>2–5</sup>, BAT holds great promise for the treatment of obesity and metabolic syndrome<sup>1</sup>. Rodent data suggest the existence of two types of brown fat cells: constitutive BAT (cBAT), which is of embryonic origin and anatomically located in the interscapular region of mice; and recruitable BAT (rBAT), which resides within white adipose tissue (WAT)<sup>6</sup> and skeletal muscle<sup>7</sup>, and has alternatively been called beige<sup>8</sup>, brite<sup>9</sup> or inducible BAT<sup>10</sup>. Bone morphogenetic proteins (BMPs) regulate the formation and thermogenic activity of BAT<sup>10–12</sup>. Here we use mouse models to provide evidence for a systemically active regulatory mechanism that controls whole-body BAT activity for thermoregulation and energy homeostasis. Genetic ablation of the type 1A BMP receptor (*Bmpr1a*) in brown adipogenic progenitor cells leads to a severe paucity of cBAT. This in turn increases sympathetic input to WAT, thereby promoting the formation of rBAT within white fat depots. This previously unknown compensatory mechanism, aimed at restoring total brown-fat-mediated thermogenic capacity in the body, is sufficient to maintain normal temperature homeostasis and resistance to diet-induced obesity. These data suggest an important physiological cross-talk between constitutive and recruitable brown fat cells. This sophisticated regulatory mechanism of body temperature may participate in the control of energy balance and metabolic disease.

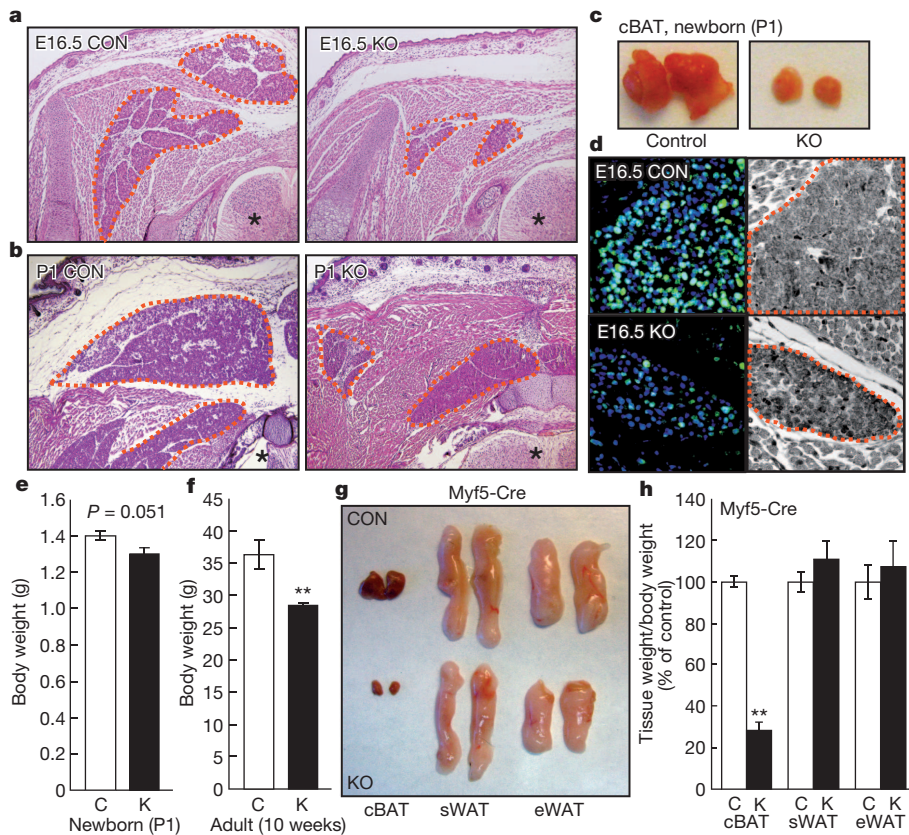
It has recently become clear that cBAT shares a common developmental ancestry with skeletal muscle<sup>13,14</sup>, whereas rBAT, localized within white fat or skeletal muscle, is derived from a non-myogenic lineage<sup>10,13</sup>. Lineage-tracing experiments have also revealed that cBAT arises from progenitors located in the embryonic dermomyotome<sup>15</sup> that express the myogenic markers Pax7 and Myf5 (refs 13, 16). Thus, we generated a mouse model lacking BMPR1A in all cells descending from the Myf5<sup>+</sup> lineage (Myf5-BMPR1A-KO). No apparent changes in morphology, proliferation or apoptosis were observed during early embryonic stages (Supplementary Figs 1 and 2). Histological evidence of reduced cBAT formation in Myf5-BMPR1A-KO mice was observed at embryonic day (E)16.5 and persisted until after birth (postnatal day (P)1; Fig. 1a–c). cBAT arises from highly proliferative fibroblasts during late gestational stages<sup>17</sup>. At E16.5, developing cBAT stains strongly for the proliferation marker Ki67, which was markedly decreased in Myf5-BMPR1A-KO embryos (Fig. 1d and Supplementary Fig. 2b, d). Apoptosis levels were unchanged in knockout mice compared to controls throughout embryogenesis (Supplementary Fig. 2c), suggesting that reduced proliferation occurring before or around E16.5 is responsible for the defective formation of cBAT in knockout animals.

Myf5-BMPR1A-KO mice were born runted and stayed smaller throughout life (Fig. 1e, f). Importantly, the reduction of cBAT mass remained highly significant in adult mice (Fig. 1g, h). Despite this, the gene-expression pattern of the residual cBAT appeared normal, apart from a moderate reduction of *Bmpr1a* gene expression (Supplementary Fig. 3). The sizes of interscapular WAT (iWAT) and retroperitoneal WAT (rWAT)—two white-fat depots that contain subpopulations of cells from the Myf5<sup>+</sup> lineage<sup>18</sup>—were also reduced in the knockout mice (Supplementary Fig. 4a). Gene expression in iWAT showed a trend towards reduced expression of BAT genes, but there were no changes in the expression of general white adipogenic genes (Supplementary Fig. 4b). Subcutaneous WAT (sWAT) and epididymal WAT (eWAT), both mostly originating from a Myf5<sup>−</sup> lineage<sup>18</sup>, were not decreased in size and expressed normal levels of all type-1 BMP receptors (Fig. 1g, h and Supplementary Fig. 3a–c). Expression of *Bmpr1a* in skeletal muscle, on the other hand, was reduced by >95% (Supplementary Fig. 3a). Upon normalization to body weight, we found that limb skeletal muscle size was unchanged, whereas the function of myogenic progenitors was altered (P.H. and E.G., unpublished data). Thus, loss of BMP signalling in Myf5-expressing cells specifically targets the formation of cBAT. During embryogenesis, MyoD<sup>+</sup> progenitors emerge after the Myf5<sup>+</sup> progenitors<sup>19</sup>. MyoD-Cre-driven *Bmpr1a* knockout mice showed completely normal development of cBAT and WAT depots (Supplementary Fig. 5), suggesting that the developmental divergence between myogenic and brown adipogenic progenitors takes place before the emergence of MyoD<sup>+</sup> progenitors, or that BMPR1A is not required for cBAT formation during this particular developmental stage.

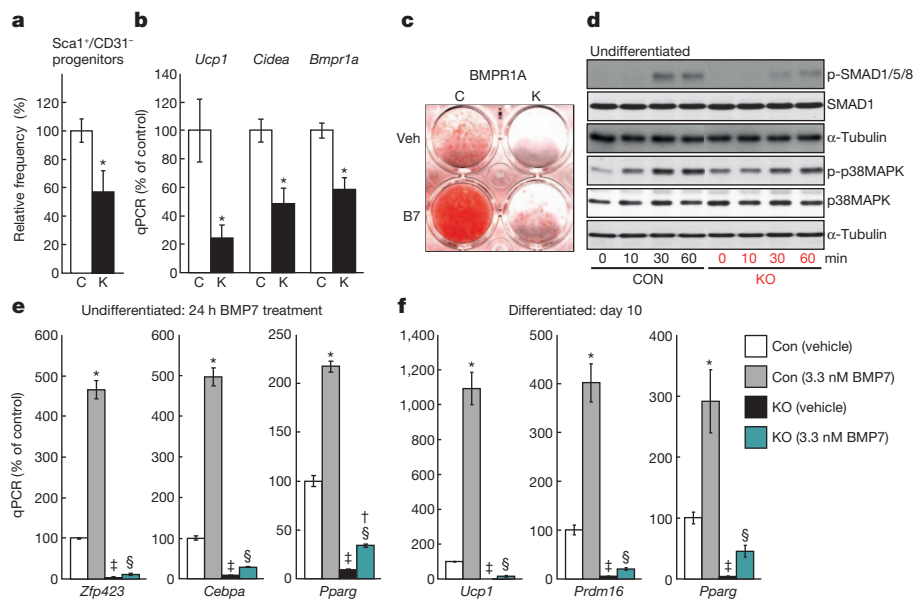
A very similar phenotype was observed in a second mouse model with conditional deletion of *Bmpr1a* in all types of adipocytes (FABP4-BMPR1A-KO). Loss of *Bmpr1a* led to a specific paucity of cBAT as well as WAT-resident rBAT (Supplementary Fig. 6). Because the FABP4-Cre driver is active at later adipogenic stages compared to Myf5-Cre, we conclude that signalling through BMPR1A is essential also for later stages of brown adipogenesis.

Next, we isolated progenitors from either Myf5<sup>−</sup> sWAT or Myf5<sup>+</sup> cBAT to test their cell-autonomous ability to produce brown adipocytes<sup>10</sup>. Cells derived from sWAT differentiated normally (Supplementary Fig. 7). By contrast, the frequency and ability of cBAT-derived Myf5<sup>+</sup>/Sca1<sup>+</sup>/CD31<sup>−</sup> (Sca1 also known as Ly6a; CD31 also known as Pecam1) progenitors to differentiate into mature brown adipocytes was significantly reduced (Fig. 2a, b). We therefore generated brown pre-adipocytes from cBAT completely lacking BMPR1A (Supplementary Fig. 8a, b). Loss of *Bmpr1a* resulted in a marked inhibition of differentiation (Fig. 2c), and reduced ability to respond to BMP7-induced phosphorylation of SMAD and p38-mitogen-activated protein kinase (p38MAPK; also known as MAPK14)<sup>20</sup>, which are downstream targets

<sup>1</sup>Section on Integrative Physiology and Metabolism, Joslin Diabetes Center, Harvard Medical School, Boston, Massachusetts 02215, USA. <sup>2</sup>Program in Genomics and Division of Genetics, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>3</sup>Division of Endocrinology and Metabolism, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai 200032, China. <sup>4</sup>Department of Biologic and Materials Sciences, School of Dentistry, University of Michigan, Ann Arbor, Michigan 48109, USA. <sup>5</sup>Harvard Stem Cell Institute, Harvard University, Cambridge, Massachusetts 02138, USA. †Present address: Research Group Adipocyte Development, German Institute of Human Nutrition, Potsdam-Rehbrücke, Nuthetal 14558, Germany.



**Figure 1 | Loss of *Bmpr1a* impairs cBAT formation by decreasing embryonic progenitor proliferation.** **a, b**, Haematoxylin and eosin staining of interscapular sections at embryonic stage E16.5 and in newborns (P1). Photographs (original magnification,  $\times 100$ ) of the same anatomical location of transversal sections were acquired for control (CON) and Myf5-BMPR1A-KO mice (KO). Asterisks indicate spinal cord, orange lines indicate brown fat. **c**, Macroscopic images of cBAT of P1 newborns. **d**, Ki67 immunofluorescence (green) of cBAT at E16.5 and co-localization with nuclear 4',6-diamidino-2-phenylindole (DAPI) stain (blue, left panels; analysis of Ki67<sup>+</sup> nuclei in Supplementary Fig. 1d), and light microscopy of the same area (right panels; original magnification,  $\times 400$ ). Orange lines indicate brown fat. **e, f**, Body weight analysis of control (C) and Myf5-BMPR1A-KO (K) mice from P1 newborns (**e**; control:  $n = 12$ ; knockout:  $n = 4$ ) and adult mice (**f**;  $n = 6$ ). **g, h**, Analysis of cBAT, sWAT and eWAT weights after normalization to body weight ( $n = 13$ ). All data are presented as mean  $\pm$  standard error of the mean (s.e.m.) Asterisks denote significant differences between genotypes:  $**P < 0.01$ .



**Figure 2 | Ablation of *Bmpr1a* in brown pre-adipocytes derived from the Myf5<sup>+</sup> lineage inhibits differentiation.** **a**, Relative frequency of Sca1<sup>+</sup>/CD31<sup>+</sup>/CD11b<sup>+</sup>/CD45<sup>+</sup> adipogenic progenitors of the Myf5-Cre(YFP)<sup>+</sup> lineage in cBAT after normalization to tissue weight ( $n = 6$ ). White bars indicate cells isolated from control mice (C), black bars indicate Myf5-BMPR1A-KO cells (K). **b**, Quantitative polymerase chain reaction (qPCR) analysis of *Ucp1*, *Cidea* and *Bmpr1a* messenger RNAs in fluorescence-activated cell sorting (FACS)-purified primary progenitors from cBAT after adipogenic differentiation ( $n = 3$ ). Asterisks denote significant differences between genotypes:  $*P < 0.05$ . **c**, Triglyceride-specific Oil-Red-O staining of differentiated control (C) or BMPR1A-deficient (K) immortalized pre-adipocyte cell lines pretreated with vehicle (Veh) or BMP7 (B7). **d**, Western blot of phospho (p)-SMAD1/5/8, basal SMAD1, corresponding  $\alpha$ -tubulin,

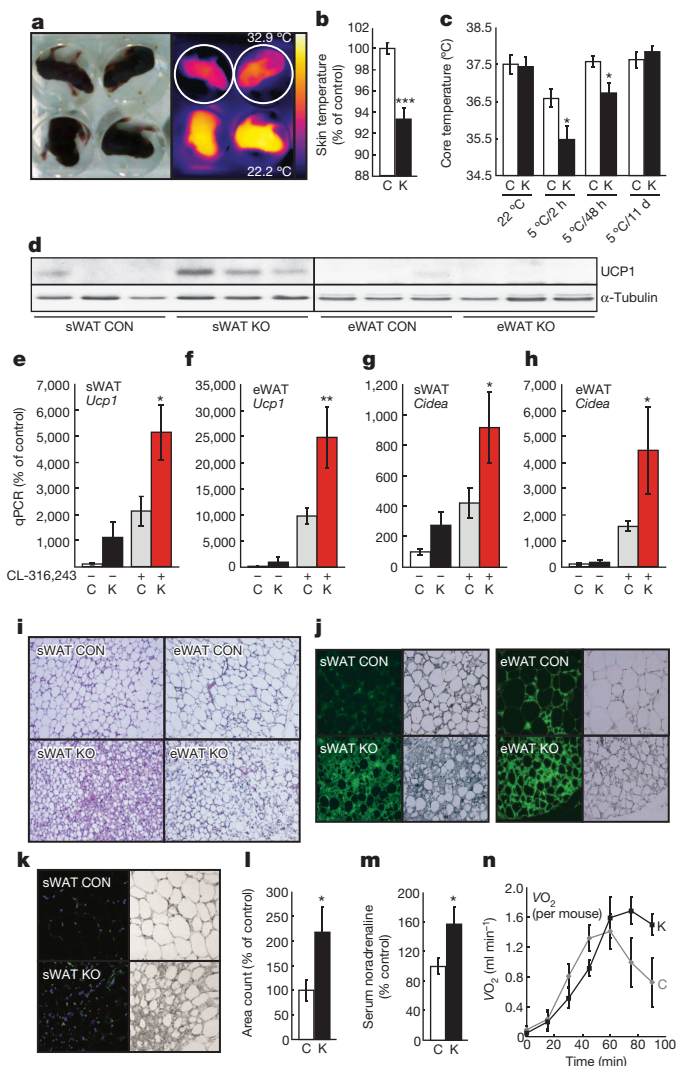
phospho-p38MAPK, basal p38MAPK, and corresponding  $\alpha$ -tubulin in undifferentiated control (CON) and pre-adipocytes lacking BMPR1A (KO) following exposure to BMP7 for 10, 30 and 60 min. Representative blots from three independent experiments are shown. **e**, qPCR of *Zfp423*, *Cebpa* and *Pparg* in undifferentiated brown pre-adipocytes after 24 h exposure to BMP7. **f**, qPCR of *Ucp1*, *Prdm16* and *Pparg* isolated from mature brown adipocytes. Pre-adipocytes were treated with BMP7 for 3 days, followed by 7-day differentiation. All experiments were performed in triplicate and are presented as mean  $\pm$  s.e.m. **e, f**, Statistically significant differences as determined by analysis of variance (ANOVA;  $P < 0.05$ ) were: control (vehicle) versus control (BMP7) (\*); KO (vehicle) versus KO (BMP7) (†); control (vehicle) versus KO (vehicle) (‡); and control (BMP7) versus KO (BMP7) (§).



of BMP signalling (Fig. 2d). This led to a concomitant decrease in the expression of key adipogenic transcription factors zinc-finger protein 423 (*Zfp423*)<sup>21</sup>, CCAAT/enhancer-binding protein  $\alpha$  (*Cebpa*) and peroxisome-proliferator-activated receptor  $\gamma$  (*Pparg*)<sup>22</sup> in undifferentiated pre-adipocytes (Fig. 2e). After adipogenic differentiation, BMPRIA-knockout cells showed severely reduced expression of BAT markers *Ucp1*, PR-domain containing 16 (*Prdm16*)<sup>13</sup> and *Pparg* (Fig. 2f), even in the presence of BMP7. Aside from BMPRIA, BMPRIIB and activin A receptor, type 1 (ACVR1) are the other two major type-1 BMP receptors<sup>20</sup>. Whereas cBAT developed normally in whole-body BMPRIIB-knockout mice (Supplementary Fig. 9a, b), deletion of *Acvr1* in the *Myf5*<sup>+</sup> lineage resulted in a severe reduction of cBAT mass (Supplementary Fig. 9c, d), indicating that both ACVR1 and BMPRIA are essential for development of cBAT. Accordingly, pre-adipocytes lacking *Acvr1* showed a somewhat milder phenotype compared to BMPRIA-deficient cells, whereas in double-knockout cells, brown adipogenesis was completely abolished (Supplementary Fig. 8c–g).

As BAT has a key role in thermoregulation, one would anticipate that the *Myf5*-BMPRIA-KO mice would have reduced body temperature as a consequence of impairment in cBAT development. Indeed, newborn *Myf5*-BMPRIA-KO mice, with their unfavourable surface-to-volume ratio, showed a significant reduction in body temperature (Fig. 3a, b). This reduction in body temperature was, surprisingly, no longer present in adult *Myf5*-BMPRIA-KO mice maintained at normal ambient temperatures of 22 °C (Fig. 3c, bars labelled 22 °C), indicating a compensatory mechanism that restored thermogenic capacity. Non-shivering thermogenesis is a critical response to prolonged cold exposure in order to maintain body temperature<sup>23</sup>. When exposed to acute and chronic cold challenges, control mice were able to quickly return to normal body temperature after 48 h of cold exposure, suggesting rapid activation of non-shivering thermogenesis via cBAT, and possibly other short-term measures such as muscle shivering, which are less pronounced in newborn mice. By contrast, *Myf5*-BMPRIA-KO mice showed a reduction in body temperature after 2 h and 48 h of cold exposure at 5 °C, presumably due to the paucity of cBAT (Fig. 3c, bars labelled 5 °C/2 h and 5 °C/48 h), as we did not observe any abnormal behaviour, such as increased shivering, under cold exposure. Despite this, body temperature in *Myf5*-BMPRIA-KO mice returned to normal after prolonged cold exposure (that is, 11 days), strongly suggesting an adaptive recruitment of rBAT to cope with lower ambient temperatures (Fig. 3c, bars labelled 5 °C/11 d). Accordingly, knockout animals showed a marked increase in UCP1 protein expression in sWAT (Fig. 3d and Supplementary Fig. 10). This browning effect could be further enhanced in sWAT, and induced in eWAT, by administration of the  $\beta$ 3-adrenergic receptor agonist CL-316,243, as indicated by significantly increased expression of BAT markers *Ucp1* and cell-death-inducing DFFA-like effector A (*Cidea*) in *Myf5*-BMPRIA-KO mice (Fig. 3e–h), as well as by increased emergence of multilocular UCP1<sup>+</sup> adipocytes in WAT (Fig. 3i, j).

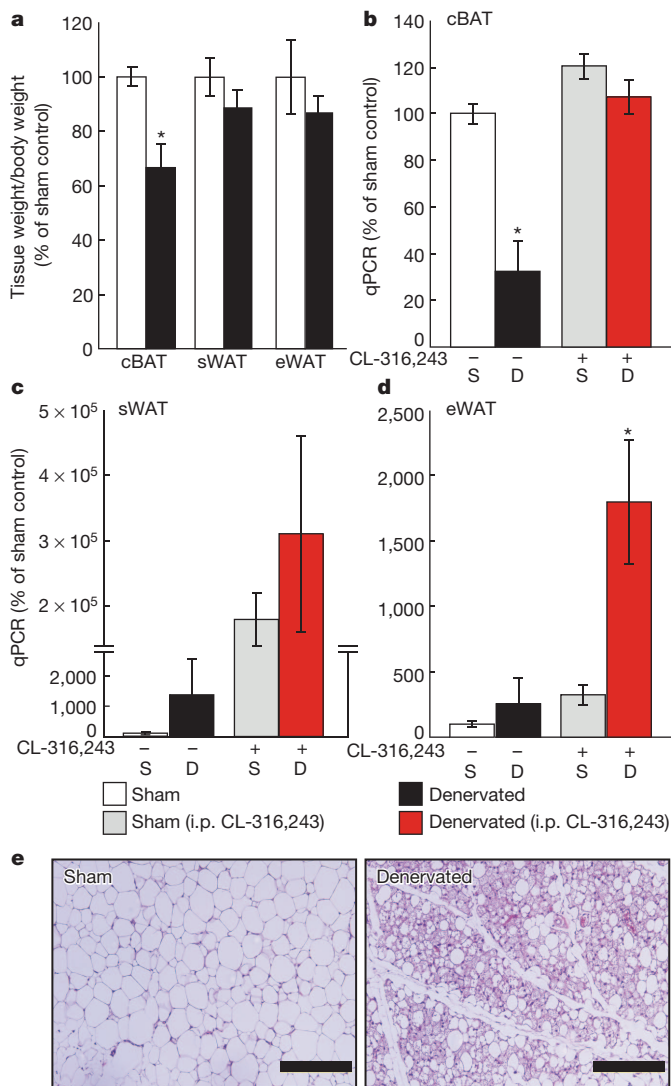
rBAT is sensitive to inductive factors, such as BMP7 (refs 10, 11), BMP8b<sup>12</sup>, fibroblast growth factor 21 (FGF21)<sup>24</sup> and the myokine irisin<sup>25</sup> (which is the processed and secreted form of the *Fndc5* gene product), among others. However, gene expression analysis revealed no changes in any of these (Supplementary Fig. 11), suggesting that the compensatory browning is not mediated by these factors. Because thermogenesis is rigorously controlled by the sympathetic nervous system<sup>26,27</sup>, we quantified sympathetic input to white fat in *Myf5*-BMPRIA-KO mice. Staining for tyrosine hydroxylase was significantly increased in sWAT of knockout mice (Fig. 3k, l). Moreover, circulating levels of noradrenaline were also significantly elevated in knockout mice, suggesting that increased sympathetic input may contribute to the browning of WAT in *Myf5*-BMPRIA-KO mice (Fig. 3m). Additionally, cold-exposed *Myf5*-BMPRIA-KO mice showed normal noradrenaline-induced thermogenic capacity (Fig. 3n and Supplementary Fig. 12), and thus possess a sufficient ability to compensate for loss of cBAT. These findings suggest that both types of brown fat may possess



**Figure 3 | Specific ablation of cBAT results in a compensatory response of increased browning by enhanced sympathetic input to WAT.**

**a**, Photographic (left) and infrared (right) images of P4–P6 newborns. Knockout pups are indicated by white circles. **b**, Quantification of average skin surface temperature from infrared images. White bars indicate control mice (C), black bars indicate *Myf5*-BMPRIA-KO mice (K; control:  $n = 12$ ; knockout:  $n = 10$ ). **c**, Body core temperature of adult control and *Myf5*-BMPRIA-KO mice. Measurements were performed at room temperature (22 °C; control:  $n = 11$ ; knockout:  $n = 10$ ), and after 2 h ( $n = 5$ ), 48 h (control:  $n = 7$ ; knockout:  $n = 6$ ) and 11 days (control:  $n = 7$ ; knockout:  $n = 6$ ) of cold exposure (5 °C). **d**, Western blot analysis of UCP1 in sWAT and eWAT. **e**, **f**, qPCR of *Ucp1* gene expression in sWAT (**e**) and eWAT (**f**). Grey bars indicate controls after administration of  $\beta$ 3-adrenergic agonist CL-316,243 (1 mg kg<sup>-1</sup> body weight) for 10 days, red bars indicate CL-316,243-injected *Myf5*-BMPRIA-KO mice ( $n = 7$ ; applies to all panels). **g**, **h**, qPCR of *Cidea* in sWAT (**g**) and eWAT (**h**). **i**, Haematoxylin and eosin staining showing morphology of sWAT and eWAT after administration of CL-316,243 (original magnification,  $\times 200$ ). **j**, UCP1 immunofluorescence of sWAT and eWAT after administration of CL-316,243 (original magnification,  $\times 400$ ). Left panels show UCP1 immunofluorescence (green), right panels show light microscope image of the same area. **k**, **l**, Tyrosine hydroxylase immunofluorescence and quantification of tyrosine-hydroxylase-positive nerve fibres in sWAT after administration of CL-316,243 (original magnification,  $\times 400$ ) ( $n = 3$ ). **m**, Serum levels of noradrenaline ( $n = 8$ ). **n**, Time course of noradrenaline-induced oxygen consumption in mice maintained at 5 °C for 8 days (grey line, control; black line, knockout;  $n = 5$ ). All data are presented as mean  $\pm$  s.e.m. Asterisks denote significant differences between genotypes: \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .





**Figure 4 | Loss of sympathetic innervation causes atrophy of cBAT and compensatory browning of white fat.** **a**, Tissue weight of cBAT, sWAT and eWAT normalized to body weight in sham-operated mice (white bars) and mice after surgical denervation of cBAT (black bars). **b–d**, qPCR of *Ucp1* in cBAT (**b**), sWAT (**c**) and eWAT (**d**) of control mice, denervated mice, controls (grey bars) or denervated (red bars) mice injected with CL-316,243. S, sham-operated mice; D, surgically denervated mice. i.p., intraperitoneal. **e**, Haematoxylin and eosin staining of sWAT from sham-operated and cBAT-denervated mice after CL-316,243 injections. Scale bar, 100  $\mu$ m. All data are presented as mean  $\pm$  s.e.m. ( $n = 4$ ). Asterisks denote significant differences between denervation and respective control groups:  $*P < 0.05$  as determined by ANOVA.

similar capacities for thermoregulation if maximally stimulated. Whereas cBAT is essential during acute cold challenges, compensatory rBAT in the knockout mice with severe paucity of cBAT has a critical role in maintaining normal body temperature, especially during long-term cold exposure. In accordance with these findings, *Myf5*-BMPRIA-KO animals were resistant to diet-induced obesity, even under obesity-promoting thermoneutrality conditions, where mice no longer require thermogenesis to maintain body temperature<sup>28</sup> (Supplementary Fig. 13).

To determine whether the effect of compensatory browning is also present in other models of cBAT atrophy and independent of genetic intervention that could also affect skeletal muscle, we surgically interrupted innervation of cBAT in wild-type mice. Denervation of cBAT resulted in a significant decrease of cBAT size and a 68% reduction of *Ucp1* expression in cBAT ( $P = 0.0029$ ; Fig. 4a, b). As in

*Myf5*-BMPRIA-KO mice, atrophy of cBAT resulted in increased recruitment of brown adipocytes in WAT (Fig. 4c–e), thus reinforcing the notion of a systemic mechanism that regulates total BAT-mediated thermogenic capacity.

It has been documented before that surgical removal of the interscapular depot of cBAT causes activation of the remaining depots of cBAT<sup>29</sup>. We used both genetic and surgically generated models of cBAT paucity to demonstrate the existence of a physiological mechanism to ensure thermoregulation by compensatory browning of WAT. This type of BAT may be more closely related to that found in adult humans<sup>8</sup>. The system inducing formation of rBAT seems to involve cBAT–brain and brain–WAT communication that is mediated, at least in part, by the sympathetic nervous system. Interestingly, obesity resistance in mice appears to be mostly related to browning of white fat, rather than adaptive thermogenesis of cBAT<sup>30</sup>, altogether suggesting that rBAT is a key contributor to metabolic health. The findings presented here suggest that any therapeutic approach involving rBAT must take into account the tight regulation of total-BAT-mediated thermogenic capacity and systemic energy metabolism at both peripheral and possibly also central levels. Targeting these mechanisms, for instance by modifying BMP signalling to regulate BAT mass and activity, could be a feasible approach to develop obesity therapies.

## METHODS SUMMARY

**Animal models.** For conditional knockouts, strains expressing Cre recombinase under the control of *Myf5*, *Myod* or fatty-acid-binding protein 4 (*Fabp4*) promoters were crossed to mice carrying floxed alleles for *Bmpr1a* or *Acrv1* (*Myf5*-Cre only). For some experiments, *Myf5*-BMPRIA-KO mice were crossed to a Rosa26-YFP reporter strain. For *Bmpr1b*, a strain with whole-body gene deletion was used. Denervation of cBAT in male, wild-type C57BL/6 mice was performed by isolating and removing a portion of about 1–2 mm from the intercostal nerve bundles innervating both lobes of the interscapular cBAT.

**Thermal imaging of skin-surface temperature.** Measurement of skin temperature was performed using a thermal imaging camera (T300 InfraRed Camera; FLIR Systems) on newborn mice at ages P4–P6, before onset of hair growth.

**Statistical analysis.** Results are expressed as mean  $\pm$  s.e.m. Differences between two groups were assessed using unpaired two-tailed Student's *t*-test or U-Mann-Whitney test. Data involving more than two groups were assessed by ANOVA.

**Full Methods** and any associated references are available in the online version of the paper.

Received 23 May 2012; accepted 29 January 2013.

Published online 13 March 2013.

- Cannon, B. & Nedergaard, J. Metabolic consequences of the presence or absence of the thermogenic capacity of brown adipose tissue in mice (and probably in humans). *Int. J. Obes. Lond.* **34** (suppl. 1), S7–S16 (2010).
- Cypess, A. M. et al. Identification and importance of brown adipose tissue in adult humans. *N. Engl. J. Med.* **360**, 1509–1517 (2009).
- Nedergaard, J., Bengtsson, T. & Cannon, B. Unexpected evidence for active brown adipose tissue in adult humans. *Am. J. Physiol. Endocrinol. Metab.* **293**, E444–E452 (2007).
- van Marken Lichtenbelt, W. D. et al. Cold-activated brown adipose tissue in healthy men. *N. Engl. J. Med.* **360**, 1500–1508 (2009).
- Virtanen, K. A. et al. Functional brown adipose tissue in healthy adults. *N. Engl. J. Med.* **360**, 1518–1525 (2009).
- Guerra, C., Koza, R. A., Yamashita, H., Walsh, K. & Kozak, L. P. Emergence of brown adipocytes in white fat in mice is under genetic control. Effects on body weight and adiposity. *J. Clin. Invest.* **102**, 412–420 (1998).
- Almind, K., Manieri, M., Sivitz, W. I., Cinti, S. & Kahn, C. R. Ectopic brown adipose tissue in muscle provides a mechanism for differences in risk of metabolic syndrome in mice. *Proc. Natl Acad. Sci. USA* **104**, 2366–2371 (2007).
- Wu, J. et al. Beige adipocytes are a distinct type of thermogenic fat cell in mouse and human. *Cell* **150**, 366–376 (2012).
- Petrovic, N. et al. Chronic peroxisome proliferator-activated receptor  $\gamma$  (PPAR $\gamma$ ) activation of epididymally derived white adipocyte cultures reveals a population of thermogenically competent, UCP1-containing adipocytes molecularly distinct from classic brown adipocytes. *J. Biol. Chem.* **285**, 7153–7164 (2010).
- Schulz, T. J. et al. Identification of inducible brown adipocyte progenitors residing in skeletal muscle and white fat. *Proc. Natl Acad. Sci. USA* **108**, 143–148 (2011).
- Tseng, Y. H. et al. New role of bone morphogenetic protein 7 in brown adipogenesis and energy expenditure. *Nature* **454**, 1000–1004 (2008).
- Whittle, A. J. et al. BMP8B increases brown adipose tissue thermogenesis through both central and peripheral actions. *Cell* **149**, 871–885 (2012).

13. Seale, P. *et al.* PRDM16 controls a brown fat/skeletal muscle switch. *Nature* **454**, 961–967 (2008).
14. Timmons, J. A. *et al.* Myogenic gene expression signature establishes that brown and white adipocytes originate from distinct cell lineages. *Proc. Natl Acad. Sci. USA* **104**, 4401–4406 (2007).
15. Atit, R. *et al.*  $\beta$ -catenin activation is necessary and sufficient to specify the dorsal dermal fate in the mouse. *Dev. Biol.* **296**, 164–176 (2006).
16. Lepper, C. & Fan, C. M. Inducible lineage tracing of Pax7-descendant cells reveals embryonic origin of adult satellite cells. *Genesis* **48**, 424–436 (2010).
17. Schmid, P., Lorenz, A., Hameister, H. & Montenarh, M. Expression of p53 during mouse embryogenesis. *Development* **113**, 857–865 (1991).
18. Sanchez-Gurmaches, J. *et al.* *PTEN* loss in the Myf5 lineage redistributes body fat and reveals subsets of white adipocytes that arise from Myf5 precursors. *Cell Metab.* **16**, 348–362 (2012).
19. Buckingham, M. E., Lyons, G. E., Ott, M. O. & Sassoon, D. A. Myogenesis in the mouse. *Ciba Found. Symp.* **165**, 111–131 (1992).
20. Kishigami, S. & Mishina, Y. BMP signaling and early embryonic patterning. *Cytokine Growth Factor Rev.* **16**, 265–278 (2005).
21. Gupta, R. K. *et al.* Transcriptional control of preadipocyte determination by Zfp423. *Nature* **464**, 619–623 (2010).
22. Farmer, S. R. Transcriptional control of adipocyte formation. *Cell Metab.* **4**, 263–273 (2006).
23. Cannon, B. & Nedergaard, J. Nonshivering thermogenesis and its adequate measurement in metabolic studies. *J. Exp. Biol.* **214**, 242–253 (2011).
24. Fisher, F. M. *et al.* FGF21 regulates PGC-1 $\alpha$  and browning of white adipose tissues in adaptive thermogenesis. *Genes Dev.* **26**, 271–281 (2012).
25. Boström, P. *et al.* A PGC1- $\alpha$ -dependent myokine that drives brown-fat-like development of white fat and thermogenesis. *Nature* **481**, 463–468 (2012).
26. Collins, S.  $\beta$ -Adrenoceptor signaling networks in adipocytes for recruiting stored fat and energy expenditure. *Front. Endocrinol.* **2**, 102 (2011).
27. Tseng, Y. H., Cypess, A. M. & Kahn, C. R. Cellular bioenergetics as a target for obesity therapy. *Nature Rev. Drug Discov.* **9**, 465–482 (2010).
28. Feldmann, H. M., Golozoubova, V., Cannon, B. & Nedergaard, J. UCP1 ablation induces obesity and abolishes diet-induced thermogenesis in mice exempt from thermal stress by living at thermoneutrality. *Cell Metab.* **9**, 203–209 (2009).
29. Rothwell, N. J. & Stock, M. J. Surgical removal of brown fat results in rapid and complete compensation by other depots. *Am. J. Physiol.* **257**, R253–R258 (1989).
30. Xue, B. *et al.* Genetic variability affects the development of brown adipocytes in white fat but not in interscapular brown fat. *J. Lipid Res.* **48**, 41–51 (2007).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported in part by National Institutes of Health (NIH) grants R01 DK077097 (Y.-H.T.), and Joslin Diabetes Center's Diabetes Research Center (DRC; P30 DK036836 from the NIDDK), a research grant from the Eli Lilly Research Foundation, and by funding from the Harvard Stem Cell Institute (to Y.-H.T.). T.J.S. was supported by the Mary K. Iacocca Foundation and the German Research Foundation (DFG; SCHU2445/1-1). P.H. was supported by a Scientist Development Grant from the American Heart Association (0730285N). K.L.T. was supported by NIH fellowships (T32 DK007260 and F32 DK091996). R.X. was supported by Project 985III-YFX0302 and NSFC81070680 from the National Natural Science Foundation of China. The authors thank Stryker Regenerative Medicine for the gift of recombinant BMP7. We acknowledge V. Kaartinen, B. Kahn, K. Lyons and P. Soriano for providing floxed *Acvr1* mice, FABP4-Cre mice, *Bmpr1b* heterozygous mice and Myf5-Cre mice, respectively. The authors thank C. R. Kahn, L. J. Goodyear, E. Kokkotou and D. Breault for comments on the manuscript. The authors wish to thank J. LaVecchio, G. Buruzula, A. Wakabayashi, A. Pinkhasov, A. Clermont, M. Mulvey, C. Cahill and G. Sankaranarayanan for technical assistance, and E. Caniano for editorial contributions.

**Author Contributions** T.J.S. and Y.-H.T. planned most of the experiments and wrote the paper. T.J.S. performed the majority of the experiments. P.H., T.L.H., L.E.M., R.X. and K.L.T. performed some of the animal and immunofluorescence experiments and/or provided research assistance. A.M.C. helped with the infrared thermography and provided valuable research materials. Y.M. and E.G. planned some of the experiments and contributed valuable research materials.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.-H.T. ([yu-hua.tseng@joslin.harvard.edu](mailto:yu-hua.tseng@joslin.harvard.edu)).

## METHODS

**Animals.** All animal procedures were approved by the Institutional Animal Use and Care Committee at Joslin Diabetes Center. Transgenic mice carrying floxed alleles for *Acrv1* (ref. 31) or *Bmpr1a* (ref. 32) were used to generate conditional gene deletion mouse models by intercrossing with either Myf5- (ref. 33) or FABP4-driven<sup>34</sup> Cre expression as indicated. For studies involving *Bmpr1b* deletion, animals with whole-body gene deletion of this receptor were used<sup>35</sup>. Myf5-Cre-expressing animals were also crossed to Rosa26-YFP reporter mice (The Jackson Laboratory) in addition to the *Bmpr1a*-floxed alleles for Myf5-lineage tracing and GFP labelling of *Bmpr1a*-deficient cells as described before<sup>10</sup>. For genotyping, DNA was isolated from tail tip biopsies by boiling in 0.5 ml 50 mM NaOH for 15 min, followed by neutralization by addition of 50  $\mu$ l 1 M Tris-base (pH 6.8) and thorough vortexing. PCR genotyping was performed using primers as listed in Supplementary Table 1. Expected band sizes for *Bmpr1a* during gel analysis were 180 bp for the wild-type allele and 230 bp for the floxed allele. For recombination analysis of *Bmpr1a* mRNA, expected band sizes were 396 and 233 bp for the intact and recombined alleles, respectively, and 178 bp for the control PCR of exons 6 to 7. For *Acrv1* genotyping, the PCR product was subsequently digested using the restriction enzyme Bgl-I (New England Biolabs) at 37 °C overnight, yielding a band at 250 bp for wild-type mice, and a double band at 90 bp and 160 bp for floxed alleles (all three bands in heterozygotes). Expected band sizes for *Bmpr1b* were 350 bp for the wild type, and 300 bp for the null allele. Expected band sizes for the Rosa26-YFP reporter mice were 600 bp for the wild-type allele and 320 bp for the mutant allele. For Myf5-Cre genotyping, bands at 600 and 400 bp for wild-type and mutant alleles, respectively, were expected. For a general Cre PCR, a single band at approximately 350 bp indicated presence of Cre cDNA in the genome. For loading control, IRS-1 primers were added for co-amplification in the same reaction (band at approximately 500 bp) to ensure proper loading with template DNA. To stimulate the browning of white adipose depots, mice were treated with daily intraperitoneal injections of 1 mg kg<sup>-1</sup> bodyweight CL-316,243 (Sigma-Aldrich) dissolved in PBS (also used for control injections) for up to 10 days. For experiments involving high-fat-diet feeding, 4- to 6-week-old animals were placed on a diet containing 45 or 60% kilocalories from fat (Research Diets). For cold exposure and thermoneutral conditions, animals were housed at 5 °C or 30 °C, respectively, for the indicated times in a controlled environmental chamber (Caron Products & Services) with free access to food and water. Body core temperature was determined by rectal probe measurements.

**Haematoxylin and eosin staining.** Sections were prepared, processed and stained as described<sup>11</sup>.

**Immunofluorescence.** Sections were deparaffinized and prepared as described<sup>11</sup>. Primary antibodies were incubated overnight at 4 °C: Ki67 (1:200, rabbit polyclonal; Abcam); UCP1 (1:50, rabbit polyclonal; AnaSpec), tyrosine hydroxylase (1:50, rabbit polyclonal; Millipore), GFP (1:200, goat polyclonal; Novus USA). After primary antibody incubation, the sections were washed and incubated with appropriate secondary antibody (Alexa Fluor-488 (green) or -594 (red); Invitrogen) at a 1:200 dilution for 10 min in the dark. After secondary antibody incubation, sections were washed with distilled water for DAPI staining (0.1  $\mu$ g ml<sup>-1</sup> in water for 5–10 min in the dark), and mounted. Sections were kept in the dark after mounting and analysed by confocal microscopy on a Zeiss LSM-410 Invert Laser Scan Microscope (Carl Zeiss MicroImaging, Thornwood, NY), or using a fluorescence microscope (Olympus BX60F-3; Olympus Corporation). Quantification of tyrosine hydroxylase was performed by using the ImageJ software. Identical conditions and settings were used for picture acquisition and analysis. A threshold was set for each image to eliminate background and to create a binary mode image. A minimum particle size of 20 pixels was used as exclusion criteria to eliminate unspecific background and for quantification of areas that stained positive for tyrosine hydroxylase. For quantification of Ki67 staining, Ki67<sup>+</sup> nuclei were counted in areas identified as BAT by microscopic inspection of morphology and comparison to published sources<sup>36</sup>, and normalized to the total number of DAPI<sup>+</sup> nuclei in the same area. For each section and animal, images from three representative areas were analysed.

**TdT-mediated dUTP nick end staining (TUNEL assay).** For detection of DNA fragmentation, sections were deparaffinized and blocked for autofluorescence as described<sup>10</sup>. Sections were blocked in 1% BSA, 0.5–1% Triton X-100 in PBS for 1 h, followed by washes in PBS. Terminal-strand labelling was performed for 1.5 h at 37 °C in TdT buffer (30 mM Tris-HCl, 140 mM Na-Cacodylate, 1 mM cobalt (II) chloride, pH 7.2) in the presence of dATP, biotinylated dUTP and terminal deoxynucleotidyl transferase (all from Roche Applied Science). The reaction was stopped by immersion in 2 $\times$  SSC buffer and subsequent incubation with Cy3-labelled streptavidin (1:200; Jackson ImmunoResearch Laboratories) for 1 h in the dark. Sections were then stained with DAPI, mounted, and analysed as described earlier.

**RNA and protein quantification.** RNA extraction, cDNA synthesis, and quantitative real-time PCR (qPCR) were performed as described before<sup>11</sup>. For qPCR

analysis, C<sub>t</sub> values <30 were used for gene expression analysis. Protein detection by western blotting was performed as described before<sup>11</sup>. Primary antibodies were incubated overnight at 4 °C: phospho-Smad-1/5/8 (1:1,000, rabbit polyclonal), SMAD1 (1:1,000, rabbit polyclonal), phospho-p38MAPK (1:1,000, rabbit polyclonal), p38MAPK (1:1,000, rabbit polyclonal) (all from Cell Signaling Technologies), UCP1 (1:500, goat polyclonal; Santa Cruz Biotechnology) and  $\alpha$ -tubulin (1:8,000, mouse monoclonal; Sigma-Aldrich). HRP-coupled secondary antibodies (Cell Signaling Technologies) were used at 1:2,000 dilutions at room temperature for 2 h followed by detection using the ECL system.

**Cell sorting.** Scat<sup>+</sup> adipocyte progenitor cells were isolated from cBAT and sWAT of Myf5-BMPRIA-KO mice and control littermates, and differentiated as described before<sup>10</sup>.

**Immortalized pre-adipocytes.** Immortalized cell lines were generated as described previously<sup>37</sup>. In brief, cBAT from individual newborn pups (P1 or P2) of homozygous floxed parents for the respective BMP receptor was collected and pre-adipocytes were isolated by enzymatic digestion. Pre-adipocytes were immortalized by infection with SV40-expressing lentivirus and subsequent selection with puromycin. Stable cell lines were then infected with adenovirus expressing either GFP (control) or a Cre::GFP fusion construct for *in vitro* recombination (Gene Transfer Vector Core, University of Iowa). Forty-eight hours after infection, GFP<sup>+</sup> cells were collected by flow cytometry and expanded for subsequent use in experiments. DNA was isolated for PCR analysis to determine full recombination of the respective receptor gene using the primers as detailed earlier.

**Cell culture.** Pre-adipocyte cell lines were cultured as described before<sup>11</sup>, except that 2% fetal bovine serum was used during differentiation. BMP7 treatments (3.3 nM) were performed for 3 days before 48 h of adipogenic induction, followed by a differentiation period of 5 days. Oil-Red-O staining was performed as described before<sup>11</sup>.

**Thermal imaging of skin-surface temperature.** Measurement of skin temperature was performed using a thermal imaging camera (T300 InfraRed Camera; FLIR Systems). Skin-surface temperature of newborns was analysed using FLIR Reporter 8.5 software (FLIR Systems). Images were acquired by placing newborn mice of the same litter (P4–P6) in 6-well cell-culture dishes, and 2–3 images of each mouse from different angles were acquired to minimize temperature variations due to different postures of the animal. Software drawing tools were used to draw a region of interest (ROI) around the entire animal, and average body surface temperature was calculated using that ROI. For each animal, an average temperature value of the temperatures from all single images was calculated. These averages were then used for statistical analysis.

**Serum parameters.** Serum levels of noradrenaline were determined using a commercially available ELISA kit and according to the manufacturer's instructions (Norepinephrine (Research) ELISA; IBL America). To stabilize the noradrenaline, 1 mM EDTA and 4 mM sodium metabisulphite were added to the serum. Serum was prepared by spinning freshly collected blood in a cooled centrifuge at 6,000g for 20 min. The clear supernatant was collected and stored at –80 °C. Samples were analysed within 12 weeks after collection. To determine circulating irisin levels, a commercially available ELISA assay kit was used according to the manufacturer's specifications (Phoenix Pharmaceuticals).

**Body composition.** Relative contents of lean and fat mass were determined using dual-energy X-ray absorptiometry (DEXA) according to the manufacturer's instructions (GE Lunar PIXImus 2, General Electric Medical Systems). Animals were anaesthetized with pentobarbital (50 mg kg<sup>-1</sup>, intraperitoneally) and placed in the scanning area to measure body composition. Relative lean and fat mass were calculated by normalizing to body weight.

**Noradrenaline-induced thermogenic capacity.** Measurement of maximum thermogenic capacity was performed as described before with some modifications<sup>12,28</sup>. Animals were maintained in the cold as described earlier for 8 days before the experiment. All measurements were performed at room temperature. Oxygen consumption by indirect calorimetry was assessed using the Comprehensive Lab Animal Monitoring System (CLAMS; Columbus Instruments) and the Oxymax for Windows software (version 4.58) for data analysis. Animals were anaesthetized with pentobarbital (80 mg kg<sup>-1</sup>, intraperitoneally), and indirect calorimetry was performed immediately for 30 min to record basal values of oxygen consumption. The sampling interval was set to 15 min to allow for stable oxygen consumption assessment throughout the experiment. After three data points were obtained, animals were briefly removed from the chamber and noradrenaline was injected subcutaneously (1 mg noradrenaline bitartrate per kg; Sigma-Aldrich), and oxygen consumption was recorded for another 90 min. To determine maximum noradrenaline-induced thermogenic capacity ( $\Delta VO_2$ ), the average value for basal oxygen consumption before noradrenaline injection was calculated and subtracted from the average value of highest noradrenaline-induced oxygen consumption ( $t_{60}$  and  $t_{75}$ ). Areas under the curve (AUC) were calculated for the curves after noradrenaline injection. Data were presented either



not normalized, that is, per animal, or after normalization to total body weight, or total fat mass (from DEXA scan).

**Denervation of cBAT.** Denervation of cBAT was performed as described before<sup>38</sup>. C57BL/6J mice (The Jackson Laboratory) aged 6 weeks were used for denervations. In brief, mice were anaesthetized and placed on a warm pad to maintain body temperature. Under a stereomicroscope, an incision was made posterior to the interscapular cBAT pad. Surrounding muscle and white fat was carefully moved to the side and the cBAT pad was turned upward to expose the five branches of the intercostal nerve bundles. Denervation was performed by isolating and cutting the nerve bundles and removing a portion of about 1–2 mm from each strand. Care was taken not to damage the adjacent blood vessels. The procedure was performed on left and right lobes of the interscapular brown fat. BAT pads, surrounding white fat, and muscle were placed in the original locations and the incision was closed using tissue adhesive glue (Vetbond, 3M Animal Care Products). The same procedure was performed for sham surgeries, except that nerve bundles were not cut. Animals were housed in single cages and monitored daily during the first week of recovery. After 10 weeks, animals were divided in four groups: (1) sham-operated mice receiving intraperitoneal vehicle injections (PBS); (2) sham-operated mice receiving injections of CL-316,243 as described earlier; (3) denervated mice receiving vehicle; or (4) CL-316,243.

**Statistical analysis.** All statistical analyses were performed using the programs Excel (Microsoft), Statview (SAS Institute) and GraphPad Prism (GraphPad

Software). Statistical analyses were performed using two-tailed Student's *t*-test or ANOVA for comparing the means of two or multiple groups, respectively. Nonparametric testing (U-Mann–Whitney test) was used where appropriate, that is, when normal distribution of sample sets was not evident. The means of two groups were considered significantly different when  $P < 0.05$ .

31. Dudas, M., Sridurongrit, S., Nagy, A., Okazaki, K. & Kaartinen, V. Craniofacial defects in mice lacking BMP type I receptor *Alk2* in neural crest cells. *Mech. Dev.* **121**, 173–182 (2004).
32. Mishina, Y., Hanks, M. C., Miura, S., Tallquist, M. D. & Behringer, R. R. Generation of *Bmpr/Alk3* conditional knockout mice. *Genesis* **32**, 69–72 (2002).
33. Tallquist, M. D., Weismann, K. E., Hellstrom, M. & Soriano, P. Early myotome specification regulates PDGFA expression and axial skeleton development. *Development* **127**, 5059–5070 (2000).
34. Abel, E. D. *et al.* Adipose-selective targeting of the *GLUT4* gene impairs insulin action in muscle and liver. *Nature* **409**, 729–733 (2001).
35. Yi, S. E., Daluiski, A., Pederson, R., Rosen, V. & Lyons, K. M. The type I BMP receptor BMPRII is required for chondrogenesis in the mouse limb. *Development* **127**, 621–630 (2000).
36. Kaufman, M. H. *The Atlas of Mouse Development* Vol. 1 (Academic, 1992).
37. Klein, J., Fasshauer, M., Klein, H. H., Benito, M. & Kahn, C. R. Novel adipocyte lines from brown fat: a model system for the study of differentiation, energy metabolism, and insulin action. *Bioessays* **24**, 382–388 (2002).
38. Puliniikunnil, T. *et al.* Adrenergic regulation of AMP-activated protein kinase in brown adipose tissue *in vivo*. *J. Biol. Chem.* **286**, 8798–8809 (2011).

# Natural RNA circles function as efficient microRNA sponges

Thomas B. Hansen<sup>1</sup>, Trine I. Jensen<sup>1</sup>, Bettina H. Clausen<sup>2</sup>, Jesper B. Bramsen<sup>1,3</sup>, Bente Finsen<sup>2</sup>, Christian K. Damgaard<sup>1</sup> & Jørgen Kjems<sup>1,3</sup>

MicroRNAs (miRNAs) are important post-transcriptional regulators of gene expression that act by direct base pairing to target sites within untranslated regions of messenger RNAs<sup>1</sup>. Recently, miRNA activity has been shown to be affected by the presence of miRNA sponge transcripts, the so-called competing endogenous RNA in humans and target mimicry in plants<sup>2–7</sup>. We previously identified a highly expressed circular RNA (circRNA) in human and mouse brain<sup>8</sup>. Here we show that this circRNA acts as a miR-7 sponge; we term this circular transcript ciRS-7 (circular RNA sponge for miR-7). ciRS-7 contains more than 70 selectively conserved miRNA target sites, and it is highly and widely associated with Argonaute (AGO) proteins in a miR-7-dependent manner. Although the circRNA is completely resistant to miRNA-mediated target destabilization, it strongly suppresses miR-7 activity, resulting in increased levels of miR-7 targets. In the mouse brain, we observe overlapping co-expression of ciRS-7 and miR-7, particularly in neocortical and hippocampal neurons, suggesting a high degree of endogenous interaction. We further show that the testis-specific circRNA, sex-determining region Y (Sry)<sup>9</sup>, serves as a miR-138 sponge, suggesting that miRNA sponge effects achieved by circRNA formation are a general phenomenon. This study serves as the first, to our knowledge, functional analysis of a naturally expressed circRNA.

To regulate mature miRNA activity, conceptual studies overexpressing miRNA target site concatamers, so-called miRNA sponges, were initially performed and shown to result in a loss of miRNA function accompanied with increased levels of endogenous targets<sup>10</sup>. Subsequently, endogenously expressed linear RNAs have been shown to sequester and inhibit miRNA activity in plants (target mimicry)<sup>2</sup> and mammals (competing endogenous RNA)<sup>3–7</sup>. Collectively, these data suggest that miRNA sponges are widespread regulators of miRNA activity in many eukaryotes. However, the miRNA sponges or miRNA-competing transcripts identified so far are expressed at low levels, contain a limited number of miRNA target sites, and, similar to miRNA targets, are themselves subject to miRNA-mediated destabilization<sup>3</sup>.

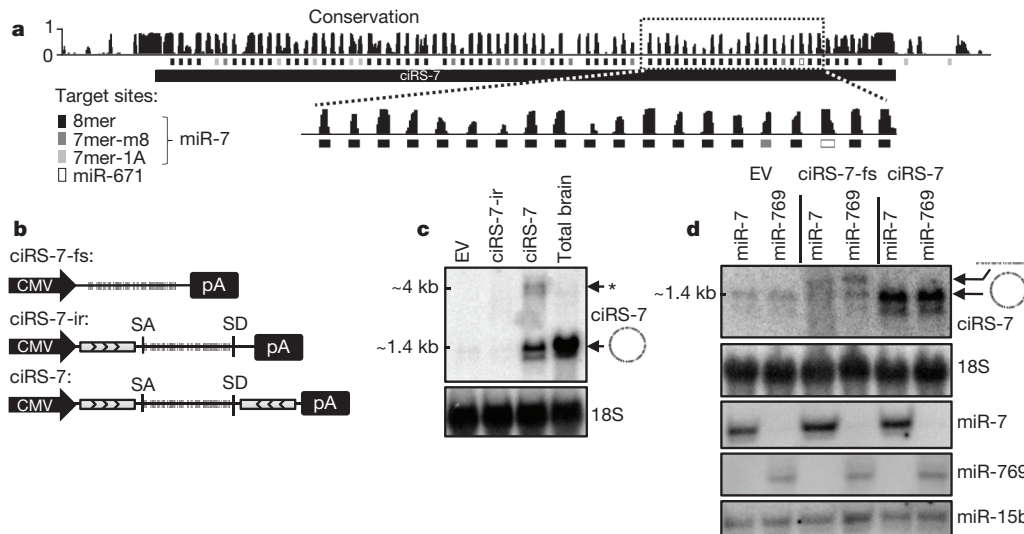
We recently identified a highly expressed endogenous circular RNA (circRNA) in human and mouse brain<sup>8</sup>, and demonstrated that the circRNA can be endonucleolytically cleaved by miR-671 in an AGO2-dependent manner; however, the function of this circRNA remained unknown. By searching for miRNA target sites in this circRNA, we identified 73 conventional seed-targets for miR-7. Plotting the target site position and nucleotide conservation, we observed that the target sites are selectively conserved compared to the adjacent regions in all eutherian mammals but not in marsupials (Fig. 1a). Although variable in size, this circRNA, ciRS-7, exists in the genome of all eutherian mammals investigated (Supplementary Fig. 1a). A computational motif analysis of human ciRS-7 using MEME<sup>11</sup> revealed a high prevalence of miR-7 seed-targets with additional 3'-end base pairing; however, they are mismatched at the central part of the duplex preventing

miRNA-mediated endocleavage<sup>12,13</sup> (Supplementary Fig. 1b). The miR-671 target site exhibits near-perfect complementarity and very little variation across species (Supplementary Fig. 1c), as well as high duplex stability compared to miR-7 (Supplementary Fig. 1d). Analysis of available online AGO2 immunoprecipitation followed by high-throughput sequencing (HITS-CLIP) data from mouse brain<sup>14</sup> revealed a high degree of AGO2 occupancy on the mouse variant of ciRS-7, with sequence reads dispersed throughout ciRS-7, and the highest read densities covering the region with the highest 8-nucleotide target site abundance (Supplementary Fig. 2a). Importantly, HITS-CLIP reads proximal to the nonlinear splice sites support the association between AGO2 and ciRS-7 (Supplementary Fig. 2b). This confirms our previous observation showing a high expression of ciRS-7 in human and mouse brain<sup>8</sup> and other studies profiling brain-specific expression of miR-7 (ref. 15).

To study the effect of ciRS-7 expression on miR-7 activity, we established a vector-based system expressing the circRNA. The biogenesis of circRNAs is at present unknown; however, previous studies have shown that exons flanked by inverted repeats induce circRNA formation *in vitro* by nonlinear splicing<sup>16</sup>, and this is thought to be essential for production of the circular Sry RNA<sup>9,17</sup>. Thus, we inserted the ciRS-7 exon along with the endogenous flanking sequence into pcDNA3 (pcDNA3-ciRS-7-ir; Fig. 1b). Subsequently, we copied part of the upstream flanking sequence and inserted it in an inverted orientation downstream (pcDNA3-ciRS-7; Fig. 1b). Transient expression of pcDNA3-ciRS-7, but not of the construct lacking the downstream inverted sequence (pcDNA3-ciRS-7-ir), resulted in the production of a circRNA species that was indistinguishable from the endogenously expressed circRNA by northern analysis (Fig. 1c) and PCR with reverse transcription (RT-PCR; Supplementary Fig. 3a). In addition, we constructed a linear and polyadenylated miR-7 sponge containing all miR-7 sponge sites but excluding the splice sites and flanking sequences (ciRS-7-fs; Fig. 1b and Supplementary Fig. 4a). Enzymatic digestion of linear RNA by tobacco acid pyrophosphatase followed by either terminator exonuclease or RNase R cleavage confirmed a circular structure of the RNA produced from ciRS-7 (Supplementary Fig. 3b, c). By co-expressing the circRNA- or the linear RNA-producing vectors with miR-7 or miR-769 expression vectors, ciRS-7 levels remained unaffected, whereas the linear construct displayed miR-7 sensitivity and was lowered by more than  $40 \pm 12\%$  (mean  $\pm$  s.e.m.) across biological replicates (Fig. 1d and Supplementary Fig. 4b), presumably owing to miRNA-mediated activation of deadenylation and subsequent exonucleolytic degradation. This suggests that the target sites in ciRS-7 do not support endocleavage by miR-7, and that ciRS-7 is resistant to the conventional miRNA destabilization of mRNA and therefore not prone to miR-7-dependent regulation.

To determine whether ciRS-7 serves as a binding platform for AGO2 and miR-7, we conducted Myc-AGO2 immunoprecipitation in HEK293 cells transiently co-expressing miR-7 or miR-769 as a control. By quantitative RT-PCR (qRT-PCR), endogenous ciRS-7

<sup>1</sup>Department of Molecular Biology and Genetics, Aarhus University, C.F. Møllers Alle 3, 8000C, Aarhus, Denmark. <sup>2</sup>Neurobiology Research, Institute of Molecular Medicine, University of Southern Denmark, J.B. Winsløvsvej 25, 5000C, Odense, Denmark. <sup>3</sup>Interdisciplinary Nanoscience Center (iNANO), Aarhus University, Gustav Wiedsvej 14, 8000C, Aarhus, Denmark.



**Figure 1 | Structure, conservation and heterologous expression of ciRS-7.** **a**, Schematic representation of the genomic locus encoding human ciRS-7 and predicted miR-7 and miR-671 target sites are shown (grey-scaled according to target site strength). Mammalian conservation scores are depicted. The horizontal black bar represents the genomic location of ciRS-7. Magnification of a ciRS-7 sub-region is shown below. 8mer, 7mer-m8 and 7mer-1A targets denote conventional miRNA seed matches (specified in Supplementary Fig. 1a). **b**, ciRS-7 expression vectors are shown. ciRS-7-fs consists of the ciRS-7 exon without flanking sequences and splice sites; ciRS-7-ir includes splice sites and the endogenous flanking genomic sequence (1 kilobase (kb) upstream and 200 base pairs (bp) downstream). In ciRS-7, part of the upstream sequence has been inverted and inserted downstream, as illustrated by the directional bars.

pull-down by AGO2 was specifically enriched more than 50-fold in miR-7-transfected cells compared to miR-769-transfected cells (Fig. 2a), strongly suggesting that miR-7 is facilitating AGO association with ciRS-7. Furthermore, using a biotin-coupled miR-7 mimic, we observed a more than sixfold enrichment of ciRS-7 in the miR-7-captured fraction compared to the negative control, biotinylated miR-138 (Fig. 2b). To strengthen the observed formation of a ternary complex between miR-7, AGO2 and ciRS-7, we devised a T7 transcript containing all putative miR-7 target sites in conjunction with or without a streptavidin-aptamer motif<sup>18</sup> (Supplementary Fig. 5a, top). As expected, by incubating the linear T7 transcripts with lysate from HEK293 cells expressing either miR-7 or control miR-769, AGO2 and miR-7 were efficiently extracted only from miR-7-containing cell lysate and in an aptamer-specific fashion (Supplementary Fig. 5a–c). This demonstrates that the miR-7 sites within ciRS-7 are able to facilitate a specific miR-7–AGO2 interaction in cell lysates. Taken together, we conclude that the widespread AGO occupancy is caused by miR-7-directed association of AGO2 proteins to the prevalent and conserved miR-7 target sites in the ciRS-7 RNA.

To evaluate the subcellular localization of miR-7 and ciRS-7, we performed RNA–fluorescence *in situ* hybridization (FISH) in HEK293 (Fig. 2c, d) and HeLa (Supplementary Fig. 6a, b) cells co-transfected with ciRS-7 and miR-7. This showed a large degree of overlap between ciRS-7 and miR-7 within distinct cellular bodies, suggesting co-localization. Immunofluorescence co-staining with the P-body-specific marker DCP1A (ref. 19) revealed that ciRS-7 co-localized with P-bodies; however, this was only observed when the circRNA was co-expressed with miR-7, suggesting that miR-7 compartmentalizes ciRS-7 to the bodies (Fig. 2d and Supplementary Fig. 6b). Endogenous ciRS-7 and miR-7 signals also overlapped in primary cells isolated from mouse brain, as determined by dual fluorophore RNA–FISH analysis (Supplementary Fig. 6c).

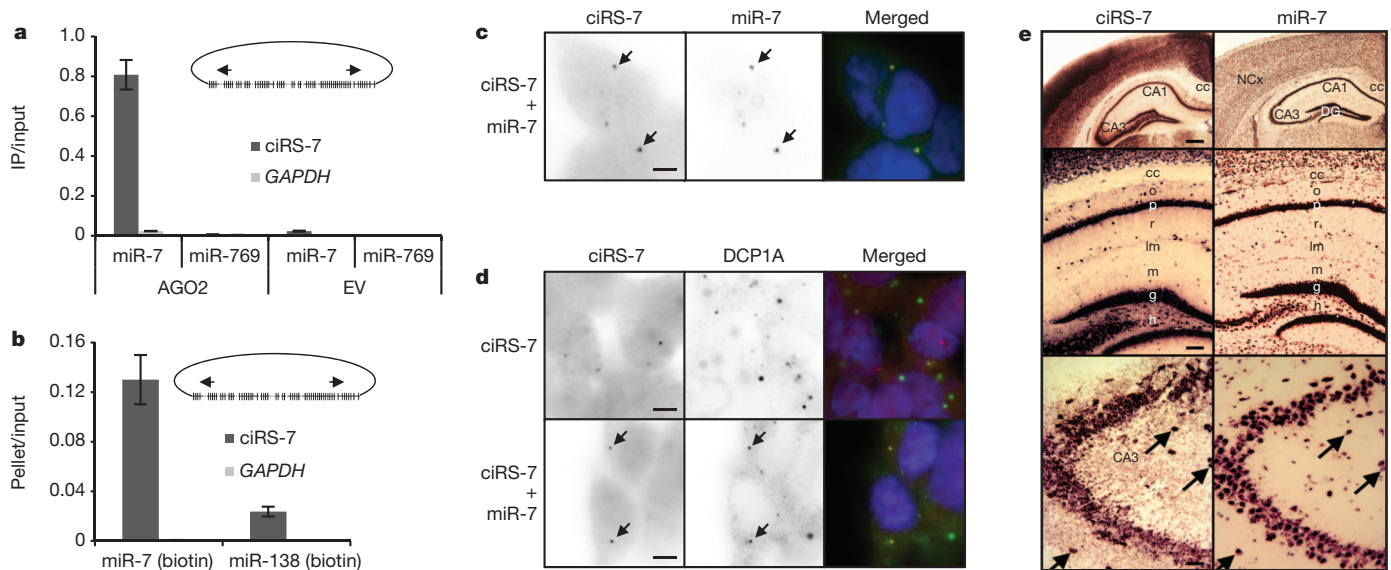
To investigate whether miR-7 and ciRS-7 are globally co-expressed in the brain, we visualized their expression in the adult mouse brain

CMV, cytomegalovirus promoter; pA, polyadenylation signal; SA, splice acceptor; SD, splice donor. **c**, Northern blot with 5 µg of RNA from HEK293 cells transfected with empty vector (EV; pcDNA3), ciRS-7-ir (pcDNA3-ciRS-7-ir), ciRS-7 (pcDNA3-ciRS-7) or total RNA from human brain. The blot was probed against ciRS-7 and 18S ribosomal RNA (loading control). Migration of circular ciRS-7 and vector-based non-circular product (asterisk) are marked. **d**, Northern blot with 10 µg (agarose northern, top two panels) or 20 µg (PAGE northern, bottom three panels) of RNA from HEK293 cells co-transfected with empty vector, ciRS-7-fs or ciRS-7, together with miR-7 or miR-769 expression vectors (pJEBB-7 or pJEBB-769, respectively). The blots were hybridized with probes as indicated. 18S rRNA and miR-15b serve as loading controls.

using alkaline phosphatase-coupled miR-7 and ciRS-7 probes. The two RNAs showed a distinct neuronal expression (Fig. 2e) compared to mouse *Gapdh* mRNA (Supplementary Fig. 7a), comprising both cortical pyramidal neurons and interneurons (Fig. 2e; see legends to Fig. 2 and Supplementary Fig. 7 for further description of brain localization). At higher magnification, it was clear that ciRS-7 exhibited dendritic translocation (Supplementary Fig. 7b). Using confocal visualization of dual fluorophore RNA–FISH (Supplementary Fig. 7c) or 5-µm thin adjacent brain slides (Fig. 2e, bottom, and Supplementary Fig. 7d, e), miR-7 and ciRS-7 displayed clear co-expression in cells. The high expression of ciRS-7 coincides with miR-7 expression in mouse brain sections and in primary cells isolated from mouse brain, which strongly suggests that miR-7 is interacting endogenously with ciRS-7.

To test the effect of ciRS-7 expression on miR-7 activity, we constructed two luciferase reporters to monitor miR-7 activity by inserting either a perfect miR-7 target (Fig. 3a) or the entire ciRS-7 sequence (Fig. 3b) into the 3′-untranslated region (UTR) of *Renilla* luciferase. Co-transfecting the luciferase reporter with miR-7 and ciRS-7 expression vectors showed a substantial reduction of knockdown potential (ranging from two- to eightfold) when ciRS-7 was co-introduced with miR-7 (Fig. 3a, b). Comparably, the non-circRNA-producing vector ciRS-7-ir had little effect on knockdown; however, this may reflect low RNA production or inefficient circularization. Notably, the expression of ciRS-7 proved more efficient than a conventional anti-miRNA approach, suggesting that the ciRS-7 molecule is indeed a highly functional miRNA sponge. Similarly, using the 3′-UTR of human *EGFR*, which is a known target for miR-7, in a luciferase reporter assay, ciRS-7 reduced the knockdown efficiency exerted by miR-7 effectively (Supplementary Fig. 8a). Next, we evaluated the effect of ciRS-7 expression on miR-7-dependent regulation of endogenous targets. To this end, a single copy of the ciRS-7 expression vector was stably inserted into HeLa Flp-In cells that otherwise have no or very little endogenous expression of ciRS-7. Analysing the expression levels of the constructed HeLa cells by northern and qRT–PCR analysis revealed





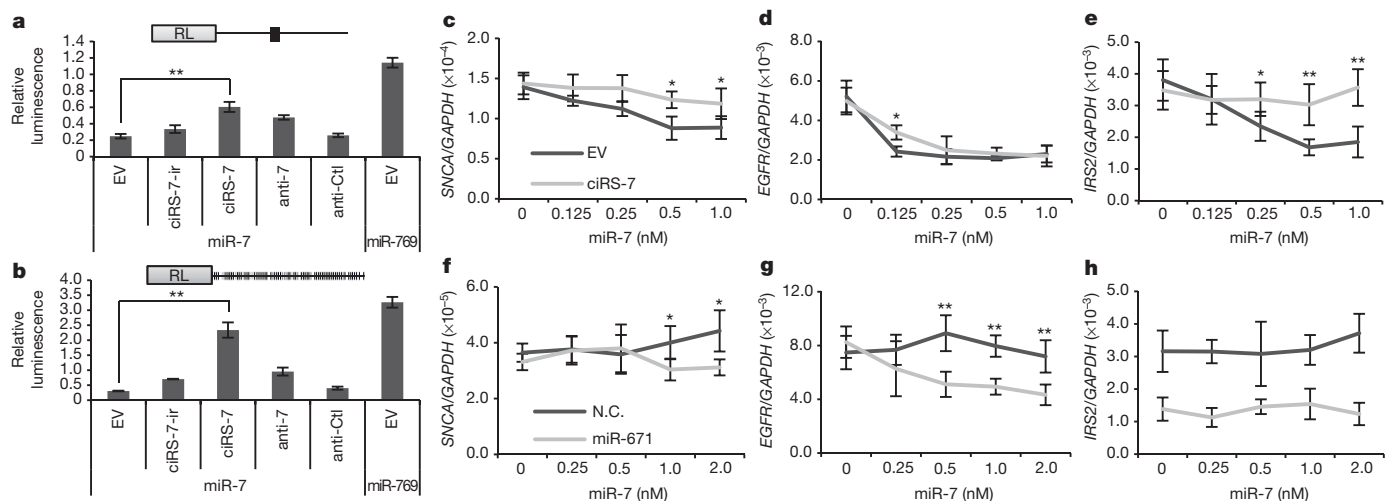
**Figure 2 | Interaction between ciRS-7 and miR-7.** **a**, Immunoprecipitation of Myc-tagged AGO2 from HEK293 cells transfected with either Myc-AGO2 plus pJEBB-7 (miR-7) or with Myc-AGO2 plus pJEBB-769 (miR-769). Empty vector serves as a Myc-AGO2 negative control. ciRS-7 and human GAPDH mRNA levels were quantified by qRT-PCR, and the relative immunoprecipitate (IP)/input ratios are plotted. Error bars represent s.d. ( $n = 4$ ). **b**, The 3'-end biotinylated miRNA duplexes were transfected into HEK293 cells. After streptavidin capture, the input and bound fractions were evaluated by qRT-PCR (analysed as in **a**). **c**, RNA-FISH on HEK293 cells co-transfected with ciRS-7 (pcDNA3-ciRS-7) and miR-7 (pCK-7). ciRS-7 and miR-7 probes were labelled with Cy5 and Cy3, respectively. Scale bar, 5  $\mu\text{m}$ . **d**, Combined immunofluorescence and RNA-FISH on ciRS-7-transfected HEK293 cells with or without miR-7 (pCK-7) co-transfection immunostained with human

DCP1A antibody and hybridized with Cy5-labelled ciRS-7 probe. Scale bars, 5  $\mu\text{m}$ . **e**, Mouse brain sections hybridized with probes specific for ciRS-7 (left) or miR-7 (right) showing intense neuronal expression throughout the neocortex (NCx), hippocampal subregions CA1 and CA3, and dentate gyrus (DG) (top and middle). Notably, miR-7 is also expressed in glial cells in the corpus callosum (cc, middle). Bottom panels show mouse brain *in situ* hybridization performed on apposing surfaces of 5- $\mu\text{m}$  thin adjacent sections. Arrows point to large neurons that based on anatomical landmarks are present in both hybridizations. g, granule cell layer; h, dentate hilus; lm, stratum lacunosum-molecular; m, stratum moleculare; o, stratum oriens CA1; p, pyramidal layer; r, stratum radiatum CA1. Scale bars, 200  $\mu\text{m}$  (top), 100  $\mu\text{m}$  (middle) and 50  $\mu\text{m}$  (bottom).

ciRS-7 levels comparable to the endogenous levels observed in HEK293 cells, although still approximately 100-fold less than the total brain levels (Supplementary Fig. 9a).

As in HEK293 cells, the stably expressed ciRS-7 in HeLa cells was sensitive towards miR-671 as shown previously<sup>8</sup>, but remained resistant to miR-7 (Supplementary Fig. 9b–d). Transfecting miR-7 in a

dose-gradient manner revealed that established miR-7 targets (such as *SNCA*<sup>20</sup>, *EGFR*<sup>21</sup> and *IRS2* (ref. 22)) responded more efficiently in the empty cell line without ciRS-7 expression (HeLa-EV) compared to ciRS-7-expressing cells (HeLa-ciRS-7) (Fig. 3c–e). In addition, when transfecting miR-671 into HeLa-ciRS-7 before introducing dose-gradient miR-7 levels, the evaluated target genes significantly regained



**Figure 3 | ciRS-7 acts as a sponge for miR-7 activity.** **a, b**, Luciferase reporter assays using reporter constructs with either a perfect miR-7 target site (**a**) or the entire ciRS-7 sequence (**b**). HEK293 cells in 12-well dishes were transfected with 0.5  $\mu\text{g}$  empty vector (pcDNA3), ciRS-7-ir (pcDNA3-ciRS-7-ir), ciRS-7 (pcDNA3-ciRS-7), 50 nM anti-miR-7 (anti-7) or 50 nM anti-miRNA control (anti-Ctl), together with 0.2  $\mu\text{g}$  miR-7 (pJEBB-7) or miR-769 (pJEBB-769) and 0.1  $\mu\text{g}$  psiCheck reporter. The relative levels of *Renilla* luminescence (RL) normalized to firefly luminescence are plotted. Error bars represent s.d. ( $n = 3$ ).

**c–e**, qRT-PCR analysis of human *SNCA* mRNA (**c**), *EGFR* mRNA (**d**) and *IRS2* mRNA (**e**) levels in HeLa-ciRS-7 (light grey line) or HeLa-EV cells (dark grey line) transfected with increasing concentrations of miR-7 as indicated. **f–h**, qRT-PCR analysis of human *SNCA* (**f**), *EGFR* (**g**) and *IRS2* (**h**) mRNA levels in HeLa-ciRS-7 cells transfected with miR-671 mimic short interfering RNA (miR-671; light grey line) or negative control (N.C.; dark grey line) siRNA 24 h before transfection with increasing concentrations of miR-7. Error bars represent s.d. ( $n = 4$ ). \* $P < 0.05$ ; \*\* $P < 0.01$ .

miR-7 sensitivity compared to control transfections (Fig. 3f–h). This suggests that miR-671 probably functions as an indirect regulator of miR-7 activity by targeting and reducing ciRS-7 levels. In line with this, the sponge effect correlates with ciRS-7 expression levels (Supplementary Fig. 8b), which after extrapolation indicates that the much higher expression of ciRS-7 in the brain exerts a stronger sponge efficiency here compared to the effect shown in HeLa-ciRS-7.

The existence of a circRNA acting as a potent miRNA sponge immediately raises the question whether circular miRNA sponges are a more general phenomenon. Indeed, by predicting miRNA target sites in the other well-known circular, testis-specific *Sry* RNA, we identified 16 putative target sites for miR-138 (Fig. 4a). To study a direct interaction between miR-138 and *Sry*, we devised a vector expressing *Sry* circRNA (Supplementary Fig. 10a). Biogenesis of circular *Sry* was verified by northern blotting and RT–PCR (Supplementary Fig. 10b, c). To validate that *Sry* circRNA can act as a miR-138 sponge, we inserted the *Sry* sequence into a luciferase reporter construct and tested the knockdown potential of miR-138 in the presence or absence of circular *Sry* RNA. Indeed, after co-expression of miR-138, we observed a fourfold reduction of relative luminescence compared to miR-769 expression, suggesting that the miR-138 target sites facilitate a miRNA effect (Fig. 4b). By co-expression of circular *Sry* RNA, the knockdown potential of miR-138 was significantly diminished (Fig. 4b), however, presumably owing to modest circular RNA biogenesis from the *Sry* expression vector (10- to 20-fold below ciRS-7, data not shown), the sponge effect here was less pronounced. Immunoprecipitation of Myc-tagged AGO2 from miR-138-transfected cells resulted in a more than tenfold enrichment of *Sry* circRNA compared to miR-769-transfected cells (Fig. 4c). Furthermore, using biotin-labelled miR-138, *Sry* circRNA was specifically captured (sixfold enrichment compared to biotin-labelled miR-7;

Fig. 4d), suggesting that *Sry* circRNA is able to interact with miR-138. Thus, we propose that the circular *Sry* RNA functions as a miR-138 sponge.

Recently, circRNA was found to be a highly prevalent RNA species in the human transcriptome<sup>23</sup>. Our observation that very abundant endogenous circRNA molecules, which are inherently resistant to exonucleolytic RNA decay, can serve as efficient miRNA sponges, adds to the growing repertoire of regulatory functions in gene expression. Considering the widespread involvement of miR-7 as a key regulator of various cancer pathways<sup>21,24</sup> and the suggested implication of miR-7 in Parkinson's disease by direct targeting of  $\alpha$ -synuclein protein expression<sup>20</sup>, ciRS-7 probably serves as a crucial factor significantly engaged in the functioning of neurons as well as a responsible candidate in neurological disorders and brain tumour development.

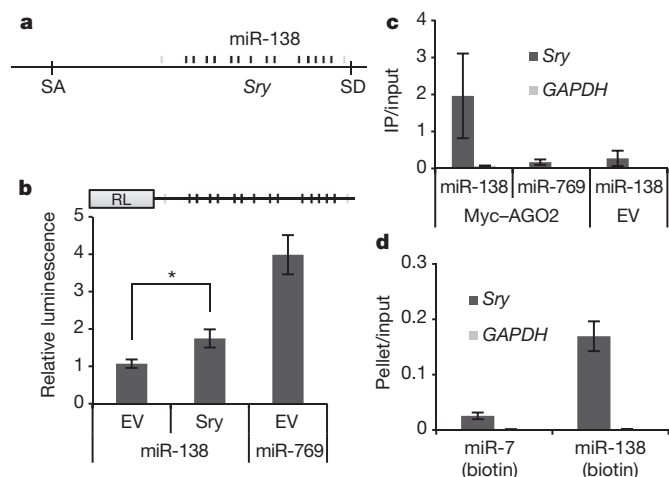
## METHODS SUMMARY

Transfections were conducted using calcium phosphate procedure or Lipofectamine 2000 (Invitrogen). RNA was isolated using TRIzol (Invitrogen) according to manufacturer's protocol. Northern blotting was conducted on 1.2% agarose or 12% PAGE, transferred to Hybond-N<sup>+</sup> membranes (GE Healthcare) and hybridized in Church buffer (see Methods). AGO immunoprecipitation was conducted as described previously<sup>25</sup>. Biotin-labelled miRNAs were transfected into HEK293 cells at a final concentration of 20 nM and treated similar to AGO immunoprecipitation. FISH was performed on paraformaldehyde-fixed cells transfected with constructs as indicated, and treated as described in Methods. Mouse brain sections (30 or 5  $\mu$ m) were made from 8–9-week-old C57BL/6 male mice. Subsequent *in situ* hybridization was performed using custom-designed alkaline phosphatase-labelled DNA and LNA probes. Luciferase assays were performed with the dual-luciferase reporter assay kit (Promega). Reverse transcription and quantitative PCR were conducted using MLV-RT (Invitrogen) and Platinum SYBR Green qPCR Supermix UDG (Invitrogen), respectively.

**Full Methods** and any associated references are available in the online version of the paper.

Received 16 August 2012; accepted 8 February 2013.

Published online 27 February 2013.



**Figure 4 | Circular *Sry* RNA interacts with miR-138.** **a**, Diagram of the mouse *Sry* locus indicating splice acceptor and splice donor sites facilitating *Sry* circularization. Perpendicular lines represent putative target sites for miR-138 (target site code as in Fig. 1a). **b**, Luciferase reporter assays with the entire *Sry* sequence. Luminescence was measured 48 h after transfection with the luciferase reporter (psiCheck-*Sry*) and expression vectors for miR-138 (pJEBB-138) or miR-769 (pJEBB-769) and with empty vector (pcDNA3; EV) or *Sry* expression vector (pcDNA3-*Sry*). The relative levels of *Renilla* luminescence normalized to firefly luminescence are plotted. Error bars represent s.d. ( $n = 3$ ). \* $P < 0.05$ . **c**, Immunoprecipitation of Myc-tagged AGO2 from HEK293 cells co-transfected with the *Sry* expression vector, Myc-AGO2 or empty vector and pJEBB-138 or pJEBB-769 as noted. Mouse *Sry* circRNA and human *GAPDH* levels were quantified by qRT–PCR and the relative immunoprecipitate/input ratios are plotted. Error bars represent s.d. ( $n = 4$ ). **d**, As described in Fig. 2b, 3'-biotinylated miR-7 or miR-138 was transfected into HEK293 cells along with the *Sry* expression vector. *Sry* circRNA and *GAPDH* levels were analysed in the input and bound fractions by qRT–PCR. Error bars represent s.d. ( $n = 4$ ).

- Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
- Franco-Zorrilla, J. M. *et al.* Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genet.* **39**, 1033–1037 (2007).
- Poliseno, L. *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038 (2010).
- Cesana, M. *et al.* A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **147**, 358–369 (2011).
- Karret, F. A. *et al.* *In vivo* identification of tumor-suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. *Cell* **147**, 382–395 (2011).
- Tay, Y. *et al.* Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell* **147**, 344–357 (2011).
- Sumazin, P. *et al.* An extensive microRNA-mediated network of RNA–RNA interactions regulates established oncogenic pathways in glioblastoma. *Cell* **147**, 370–381 (2011).
- Hansen, T. B. *et al.* miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* **30**, 4414–4422 (2011).
- Capel, B. *et al.* Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. *Cell* **73**, 1019–1030 (1993).
- Ebert, M. S., Neilson, J. R. & Sharp, P. A. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nature Methods* **4**, 721–726 (2007).
- Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36 (1994).
- Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W. & Tuschl, T. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.* **20**, 6877–6888 (2001).
- Hutvagner, G. & Zamore, P. D. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297**, 2056–2060 (2002).
- Chi, S. W., Zang, J. B., Mele, A. & Darnell, R. B. Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps. *Nature* **460**, 479–486 (2009).
- Bak, M. *et al.* MicroRNA expression in the adult mouse central nervous system. *RNA* **14**, 432–444 (2008).
- Pasman, Z., Been, M. D. & Garcia-Blanco, M. A. Exon circularization in mammalian nuclear extracts. *RNA* **2**, 603–610 (1996).
- Dubin, R. A., Kazmi, M. A. & Ostrer, H. Inverted repeats are necessary for circularization of the mouse testis *Sry* transcript. *Gene* **167**, 245–248 (1995).
- Iio, H., Loiselet, D., Haystead, T. A. & Macara, I. G. Efficient detection of RNA–protein interactions using tethered RNAs. *Nucleic Acids Res.* **39**, e53 (2011).

19. Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nature Cell Biol.* **7**, 719–723 (2005).
20. Junn, E. *et al.* Repression of  $\alpha$ -synuclein expression and toxicity by microRNA-7. *Proc. Natl Acad. Sci. USA* **106**, 13052–13057 (2009).
21. Kefas, B. *et al.* microRNA-7 inhibits the epidermal growth factor receptor and the Akt pathway and is down-regulated in glioblastoma. *Cancer Res.* **68**, 3566–3572 (2008).
22. Jiang, L. *et al.* MicroRNA-7 targets IGF1R (insulin-like growth factor 1 receptor) in tongue squamous cell carcinoma cells. *Biochem. J.* **432**, 199–205 (2010).
23. Salzman, J., Gawad, C., Wang, P. L., Lacayo, N. & Brown, P. O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS ONE* **7**, e30733 (2012).
24. Reddy, S. D., Ohshiro, K., Rayala, S. K. & Kumar, R. MicroRNA-7, a homeobox D10 target, inhibits p21-activated kinase 1 and regulates its functions. *Cancer Res.* **68**, 8195–8200 (2008).
25. Hendrickson, D. G., Hogan, D. J., Herschlag, D., Ferrell, J. E. & Brown, P. O. Systematic identification of mRNAs recruited to Argonaute 2 by specific microRNAs and corresponding changes in transcript abundance. *PLoS ONE* **3**, e2126 (2008).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank C. Bus and R. Rosendahl for technical assistance and R. M. Zadegan for art work. We also thank the G. Hannon laboratory for providing us with the Myc-tagged AGO2 expression vector, J. Lykke-Andersen for the DCP1A antibody, and K. L. Lambertsen and C. U. von Linstow for their assistance with confocal microscopy. This work was supported by the SIROCCO EU consortium, the Lundbeck Foundation, and the Danish Council for Independent Research - Natural Sciences. T.B.H. and B.H.C. were supported by the Lundbeck Foundation.

**Author Contributions** T.B.H. conceived the project, designed the experiments and drafted the manuscript. T.B.H., T.I.J. and C.K.D. performed the experiments. J.B.B. assisted experimentally and intellectually. B.H.C. performed the brain *in situ* hybridizations. B.F., C.K.D. and J.K. supervised the project and revised the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.K. (jk@mb.au.dk) or T.B.H. (tbh@mb.au.dk).



## METHODS

**Constructs and transfections.** HEK293 Flp-In T-Rex (Invitrogen) and HeLa Flp-In (Invitrogen) cells were maintained under standard culture conditions. Transfections were conducted with either calcium phosphate using standard procedures or Lipofectamine 2000 (Invitrogen) according to supplier's protocol. The construction of pJEBB-769 has been published previously<sup>8</sup>. pJEBB-7 and pJEBB-138 were constructed similarly; PCR amplification of the miR-7-1 locus was performed using miR-7 forward/reverse primers, or the miR-138 locus (in mouse) using miR-138 forward/reverse primers, and the fragments were inserted into a NotI- and SalI-digested enhanced green fluorescent protein (eGFP) expression cassette (pJEBB). pCK-7 was generated by digesting pJEBB-7 with NotI and XbaI and inserting the resulting fragment into NotI- and XbaI-digested pCK.

pcDNA3-ciRS-7-ir was constructed by PCR amplifying the circRNA locus, including 1 kb upstream and 200 bp downstream to the nonlinear splice sites, using ciRS-7 forward/reverse primers. The PCR fragment was inserted into HindIII- and NotI-digested pcDNA3 (Invitrogen). An ~800-bp DNA stretch upstream of the splice acceptor was amplified using ciRS-7-ir forward/reverse primers and inserted downstream in the reverse orientation in an XhoI-digested pcDNA3-ciRS-7-ir, thus generating pcDNA3-ciRS-7. For the generation of pcDNA3-ciRS-fs, PCR amplification of the ciRS-7 exon using ciRS-7-fs forward/reverse primers was inserted into pcDNA3 using HindIII and XhoI. To establish HeLa cells stably expressing ciRS-7, ciRS-7 was copy-pasted into pcDNA5 using HindIII and ApaI, and transfected into HeLa Flp-In cells (Invitrogen) along with pOG44. Forty-eight hours after transfections, the cell media was supplemented with hygromycin B according to manufacturer's protocol.

pcDNA3-Sry was constructed by PCR amplifying the Sry locus from male mouse DNA using Sry forward/reverse primers, and inserting the amplicon into pcDNA3 using BamHI and NotI. Subsequently, SalI-digested CMV promoter from pcDNA5 was inserted into the XhoI site in an inverse orientation.

Luciferase reporter vectors were constructed by inserting annealed primers (miR-7-psiCheck forward/reverse) or PCR amplicons (produced by ciRS-7-psiCheck forward/reverse, Sry-psiCheck forward/reverse or EGFR-psiCheck forward/reverse) into the psiCheck-2 vector.

Transfection of miRNA mimics (miR-7-5p and miR-671-5p, Applied Biosystems) was conducted using Lipofectamine 2000 (Invitrogen). In case of dose-gradient transfections, a total concentration of 10 nM was used with varying ratios between miR-7 and negative control oligonucleotide, as indicated.

**Northern blotting.** For miRNA, 20–30 µg whole cell RNA was loaded on a 12% denaturing PAGE gel. RNA was transferred to Amersham Hybond-N+ membranes (GE Healthcare). The membranes were hybridized with [<sup>32</sup>P]-labelled DNA oligonucleotides (listed in Supplementary Table 1) in Church buffer (0.5 M NaPO<sub>4</sub>, 7% SDS, 1 mM EDTA, 1% BSA, pH 7.5) at 37 °C and washed in 2× SSC (300 mM NaCl, 30 mM Na-citrate, pH 7.0) with 0.1% SDS at room temperature. The membranes were exposed on phosphorimager screens and analysed using Quantity One or Image Lab software (Bio Rad).

For circRNA, agarose northern blot was performed with 5–10 µg RNA separated in 1.2% agarose. Subsequent hybridization and wash was carried out at 55 °C and 50 °C, respectively, otherwise conducted as described above.

**AGO immunoprecipitation.** Myc-tagged AGO2 was co-transfected with either pJEBB-7 or pJEBB-769 in p10 dishes. Forty-eight hours after transfection, AGO immunoprecipitation was performed as described previously<sup>25</sup>. In brief, cells were lysed in 150 mM KCl, 25 mM Tris-HCl, pH 7.4, 5 mM EDTA, 0.5% Triton X-100 and 5 mM dithiothreitol (DTT) supplemented with Ribolock (Fermentas) and proteinase inhibitor cocktail (Roche). The lysate was mixed with Myc antibody (Abcam, ab9106) coupled sepharose beads (GE Healthcare) and left under rotation for 4 h at 4 °C. Beads were subsequently washed five times in lysis buffer and the RNA was extracted using TRIzol reagent (Invitrogen).

**Streptavidin aptamer.** T7 transcription was performed on PCR products containing the entire circRNA sequence with or without the inclusion of a streptavidin aptamer sequence using T7 MEGAscript (Ambion). The streptavidin-based capturing of RNA with associated proteins was conducted as previously published<sup>18</sup>. In brief, the T7 transcript was renatured and mixed with streptavidin beads (Dynabeads MyOne Streptavidin C1, Invitrogen) for 1 h at 4 °C and washed twice in lysis buffer (10 mM HEPES, pH 7.0, 200 mM NaCl, 1% Triton X-100, 10 mM MgCl<sub>2</sub> and 1 mM DTT). Cells transfected 48 h before collection with either pJEBB-7 or pJEBB-769 were lysed and added to the RNA-coupled beads using half a confluent p10 dish per streptavidin capture. After rotating for 4 h at 4 °C, the beads were washed five times and RNA and protein were extracted and analysed by northern blot (probes listed in Supplementary Table 1) and by western blot using anti-AGO2 (Abcam, ab57113) and anti-HuR (Santa Cruz, SC-5261) as a loading control.

**Biotin-coupled miRNA capture.** Biotin-labelled miRNAs were transfected into HEK293 cells at a final concentration of 20 nM. Forty-eight hours after transfection, cells were collected and treated as described above.

**Immunofluorescence and RNA-FISH.** Cells were seeded on glass coverslips in 12-well plates and transfected with ciRS-7 and pCK-7 using calcium phosphate (HEK293 cells) or Lipofectamine 2000 (HeLa cells). Forty-eight hours after transfection, cells were washed in PBS and fixed in 4% paraformaldehyde for 15 min and permeabilized overnight in 70% ethanol. Cells were then washed twice in PBS containing 5 mM MgCl<sub>2</sub> (PBSM), and rehydrated for 10 min in 50% formamide and 2× SSC. In case of immunofluorescence, cells were blocked with 3% BSA in PBSM for 1 h followed by incubation with primary antibody (human DCP1A (ref. 26)) in 3%BSA/PBSM supplemented with RNaseOUT (Fermentas) at 37 °C. After washing three times in PBSM, cells were incubated with secondary antibody (Alexa Fluor 488, A21206, Invitrogen). Cells were then washed three times before proceeding with RNA-FISH.

For FISH, cells were incubated at 37 °C in a solution containing 50% formamide, 2× SSC, 0.25 mg ml<sup>-1</sup> *Escherichia coli* transfer RNA, 0.25 mg ml<sup>-1</sup> salmon sperm DNA (Invitrogen), 2.5 mg ml<sup>-1</sup> BSA (Roche), and 0.5 ng ml<sup>-1</sup> fluorescently labelled ciRS-7 (Supplementary Table 1) and miR-7 (Exiqon) probes. After 3 h, cells were washed twice for 20 min at 37 °C in 50% formamide and 2× SSC, followed by four 5-min washes in PBS (the penultimate wash containing 4',6'-diamidino-2-phenylindole (DAPI)) and an additional brief wash in nuclease-free water. Cells were mounted in ProLong Gold (Invitrogen) and left overnight at room temperature.

**Mice for the *in vivo* study.** This study was performed in accordance with guidelines approved by the Danish Animal Ethical Committee (no. 2011/561-1950), using 8–9-week-old C57BL/6 male mice (*n* = 5) (Taconic). The mice were euthanized by cervical dislocation, the brains frozen in gaseous CO<sub>2</sub> and processed into 30-µm or 5-µm coronal cryostat sections. For paraformaldehyde-fixed tissue, brains were fixed and processed into 16-µm sections for histological analysis as described previously<sup>27</sup>.

***In situ* hybridization.** *In situ* hybridization was performed as previously described<sup>28</sup> using custom designed alkaline phosphatase-labelled DNA and LNA probes (Supplementary Table 1) recognizing mouse *Gapdh*, ciRS-7 and miR-7 (DNA Technology A/S). Double fluorescence *in situ* hybridization was performed on fixed tissue. Sections were rinsed for 10 min in TBS buffer containing 0.5% Triton X-100 (ref. 27) and incubated overnight with Maxima Probe/master mix (Thermo Scientific) containing Cy3-labelled ciRS-7 and Cy5-labelled miR-7 probes. Development was arrested by rinsing the sections at 37 °C in tap water and distilled water, before mounting with Aqueatex (Merck). The specificity of the signal was tested by hybridizing parallel sections with: (1) 100-fold excess of unlabelled probe, (2) pre-treatment with RNase A (Pharmacia Biotech), (3) buffer alone, or (4) Cy5-labelled miR-449 probe. All controls were devoid of signal.

Estimation of the percentage of ciRS-7- and miR-7-co-expressing neurons was performed in layer IV–V of the cingulate gyrus. ciRS-7- and miR-7-expressing neurons were counted on digitalized images of pairs of adjacent 5-µm thick sections. Per section, 75 randomly selected ciRS-7-expressing neurons were counted and analysed for co-expression of miR-7. Only ciRS-7-expressing neurons with a clearly identifiable nucleus were analysed. For control, a comparable number of circular regions devoid of ciRS-7 was analysed for miR-7 expression.

**Luciferase reporter assay.** HEK293 cells seeded in 12-well plates and collected 48 h after transfection using the dual-luciferase reporter assay kit (Promega). In case of anti-miRNA experiments, anti-miRNA inhibitor (Applied Biosystems) was used at a final concentration of 50 nM. Luminescence was measured on a BMG FLUOstar luminometer (BMG labtech).

**Tobacco acid pyrophosphatase exonuclease and RNase R treatment.** In case of tobacco acid pyrophosphatase exonuclease, 4 µg of RNA was incubated with 10 U of tobacco acid pyrophosphatase (Epicentre) at 37 °C for 1 h. After incubation, the RNA was ethanol precipitated and incubated with 1 U of Terminator 5'-dependent exonuclease (Epicentre). In case of RNase R, 4 µg of RNA was incubated with 20 U of RNase R (Epicentre) at 30 °C for 10 min. In both cases, half the RNA was visualized by northern blotting, the other half was subjected to qRT-PCR as described below.

**qRT-PCR.** Total RNA was extracted from cells in culture using TRIzol reagent (Invitrogen) according to standard procedures. Complementary DNA for qRT-PCR expression analyses was synthesized from total RNA by MLV-RT (Invitrogen) according to the supplied protocol using random hexamer primer. Quantitative PCR was performed with Platinum SYBR Green qPCR Supermix UDG (Invitrogen) on either MxPro3000 cyclor (Stratagene) or LightCycler 480 (Roche) according to standard procedures, using primers listed in Supplementary

Table 1. RNA levels were normalized to human *GAPDH*. In case of miRNA Taqman qPCR, miR-7, miR-769 and RNU48 assays (Applied Biosystems) were used adhering to manufacturer's protocol. All quantitative PCR reactions were conducted in quadruplicates.

In case of 3'-rapid amplification of cDNA ends (RACE), cDNA synthesis was primed by oligo-dTN. PCR was subsequently conducted using the 3'-RACE reverse/ciRS-7 (2) forward primer set.

**Statistical analysis.** Statistical significance was determined by the standard two-tailed Student's *t*-test.

26. Lykke-Andersen, J. & Wagner, E. Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. *Genes Dev.* **19**, 351–361 (2005).
27. Clausen, B. H. *et al.* Interleukin-1 $\beta$  and tumor necrosis factor- $\alpha$  are expressed by different subsets of microglia and macrophages after ischemic stroke in mice. *J. Neuroinflammation* **5**, 46 (2008).
28. Clausen, B. H., Lambertsen, K. L. & Finsen, B. Glyceraldehyde-3-phosphate dehydrogenase versus toluidine blue as a marker for infarct volume estimation following permanent middle cerebral artery occlusion in mice. *Exp. Brain Res.* **175**, 60–67 (2006).

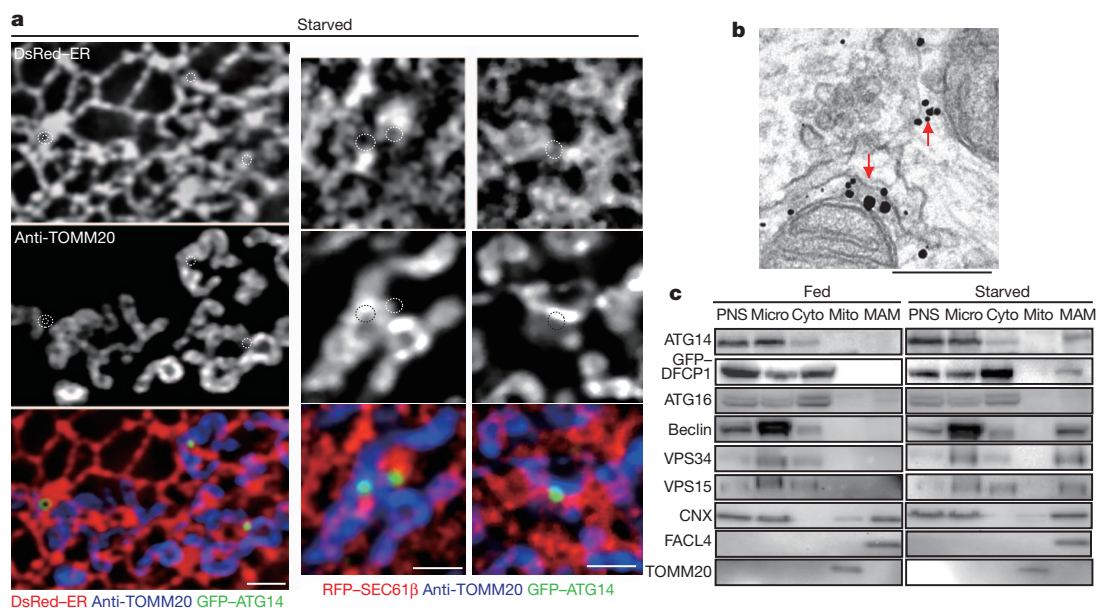
# Autophagosomes form at ER–mitochondria contact sites

Maho Hamasaki<sup>1,2\*</sup>, Nobumichi Furuta<sup>3\*</sup>, Atsushi Matsuda<sup>4,5</sup>, Akiko Nezu<sup>1,2</sup>, Akitsugu Yamamoto<sup>6</sup>, Naonobu Fujita<sup>1,2</sup>, Hiroko Oomori<sup>7</sup>, Takeshi Noda<sup>1,2</sup>, Tokuko Haraguchi<sup>4,5</sup>, Yasushi Hiraoka<sup>4,5</sup>, Atsuo Amano<sup>3,8</sup> & Tamotsu Yoshimori<sup>1,2</sup>

Autophagy is a tightly regulated intracellular bulk degradation/recycling system that has fundamental roles in cellular homeostasis<sup>1</sup>. Autophagy is initiated by isolation membranes, which form and elongate as they engulf portions of the cytoplasm and organelles. Eventually isolation membranes close to form double membrane-bound autophagosomes and fuse with lysosomes to degrade their contents. The physiological role of autophagy has been determined since its discovery, but the origin of autophagosomal membranes has remained unclear. At present, there is much controversy about the organelle from which the membranes originate—the endoplasmic reticulum (ER), mitochondria and plasma membrane<sup>1,2</sup>. Here we show that autophagosomes form at the ER–mitochondria contact site in mammalian cells. Imaging data reveal that the pre-autophagosome/autophagosome marker ATG14 (also known as ATG14L) relocates to the ER–mitochondria contact site after starvation, and the autophagosome-formation marker ATG5 also localizes at the site until formation is complete. Subcellular fractionation showed that ATG14 co-fractionates in the mitochondria-associated ER membrane<sup>3–5</sup> fraction under starvation conditions. Disruption of

the ER–mitochondria contact site prevents the formation of ATG14 puncta. The ER-resident SNARE protein syntaxin 17 (STX17) binds ATG14 and recruits it to the ER–mitochondria contact site. These results provide new insight into organelle biogenesis by demonstrating that the ER–mitochondria contact site is important in autophagosome formation.

ATG14 is a unique subunit of the autophagy-specific phosphatidylinositol-3-OH kinase (PI(3)K) complex involved in autophagosome formation<sup>6,7</sup>. The ATG14 complex localizes to the isolation membrane, the cytoplasm and the ER. The anchoring of ATG14 to the ER is essential for autophagosome formation: under fed conditions, the ATG14 complex seems to diffuse within the ER membrane; after induction of autophagy by starvation, the complex assembles at several specific points before autophagosome formation<sup>7</sup>. Three-dimensional electron tomographic studies have shown that the isolation membrane is ‘sandwiched’ between two ER cisternae, one of which is connected to the isolation membrane by a narrow extension<sup>8,9</sup>. Taken together with a previous study showing that the ER-associated membrane known as the omegasome is the formation site<sup>10</sup>, it is thought that the ER is the



**Figure 1 | ATG14 assembles at the ER–mitochondria contact site.** **a**, COS7 cells co-transfected with GFP–ATG14 and DsRed–ER or RFP–SEC61β (an ER marker) were immunostained for the mitochondria marker TOMM20. Scale bars, 2 μm. **b**, An immunogold electron micrograph of starved HeLa cells

expressing GFP–ATG14 probed with anti-GFP antibodies (red arrows). Scale bar, 200 nm. **c**, Immunoblots of subcellular fractions isolated from HEK293 cells stably expressing GFP–DFCP1. Cyto, cytosol; micro, microsome; mito, mitochondria; PNS, post-nuclear supernatant.

<sup>1</sup>Department of Genetics, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan. <sup>2</sup>Laboratory of Intracellular Membrane Dynamics, Graduate School of Frontier Biosciences, Osaka University, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan. <sup>3</sup>Department of Oral Frontier Biology, Center for Oral Frontier Science, Osaka University Graduate School of Dentistry, 1-8 Yamadaoka, Suita, Osaka 565-0871, Japan. <sup>4</sup>National Institute of Information and Communications Technology, 588-2 Iwaoka, Iwaoka-cho, Nishi-ku, Kobe 651-2492, Japan. <sup>5</sup>Nuclear Dynamics Group, Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan. <sup>6</sup>Department of Cell Biology, Faculty of Bio-Science, Nagahama Institute of Bio-Science and Technology, 1266 Tamura-cho, Nagahama, Shiga 526-0829, Japan. <sup>7</sup>Research Institute for Microbial Diseases, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565-0871, Japan. <sup>8</sup>Department of Preventive Dentistry, Osaka University Graduate School of Dentistry, 1-8 Yamadaoka, Suita, Osaka 565-0871, Japan.

\*These authors contributed equally to this work.



**Table 1 | Quantification of GFP-ATG14 localization**

Localization of ATG14 dots	<i>n</i> = 380
Adjacent to or on both ER and mitochondria	201 (52.9%)
Adjacent to or on ER only	177 (46.5%)
Adjacent to or on mitochondria only	1 (0.3%)
On neither organelle	1 (0.3%)

*n* denotes the number of ATG14 dots (described in Fig. 1a).

primary platform for autophagosome formation. Because the ER and mitochondria contact each other, we examine the relationship between autophagosome formation sites and the ER-mitochondria contact sites by triple-colour imaging. We show that ATG14 assembled exclusively at the ER-mitochondria contact site under starvation conditions (Fig. 1a, Table 1 and Supplementary Fig. 2a). Using immunoelectron microscopy, we observed that green fluorescent protein (GFP)-tagged ATG14 accumulated at ER-mitochondria contact sites under starved conditions (Fig. 1b and Supplementary Fig. 3a, b), and confirmed these observations in cells expressing Myc-ATG14 (Supplementary Fig. 3c, d). Our results support the idea that autophagosome formation takes place at the ER-mitochondria contact site during starvation.

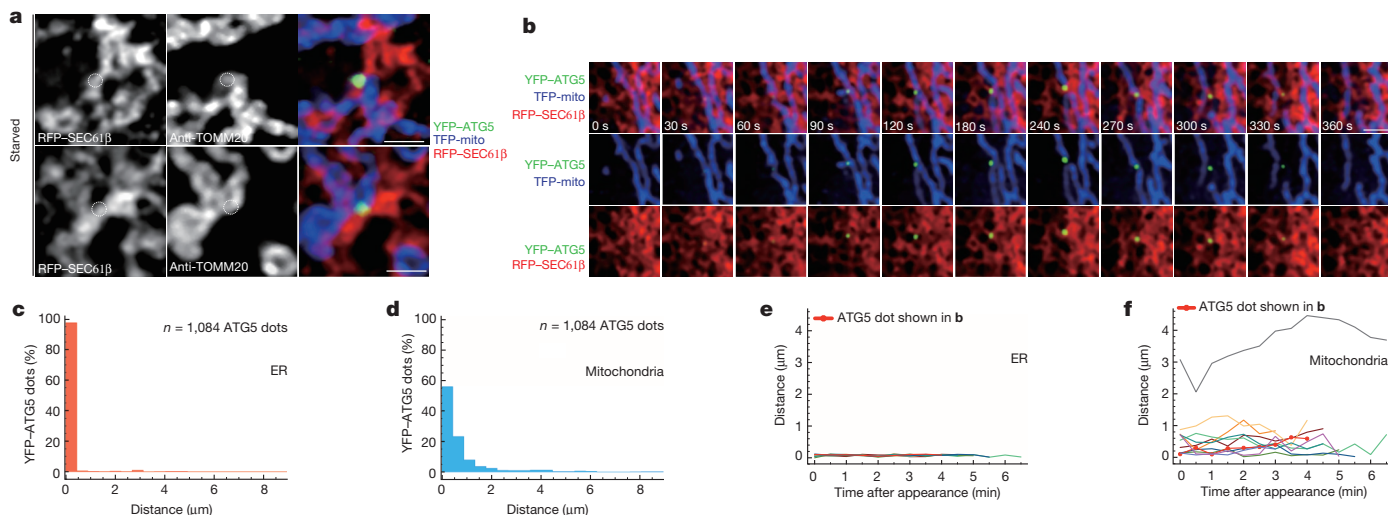
We recovered the mitochondria-associated ER membrane (MAM)<sup>3</sup>, which includes the contact sites, by Percoll centrifugation, identifying fractions using two MAM markers, the Ca<sup>2+</sup>-binding chaperone calnexin (CNX) and fatty acid CoA ligase 4 (FACL4), as well as mitochondrial translocase of outer mitochondrial membrane 20 (TOMM20)<sup>11–13</sup>. Under fed conditions, endogenous ATG14 was not present in the MAM fractions (Fig. 1c); under starved conditions, however, the localization of ATG14 shifted markedly to the MAM. The localization of other components of the PI(3)K complex shifted in a similar manner to that of ATG14 (Fig. 1c). After starvation, double FYVE domain-containing protein 1 (DFCP1) translocated to the omega-some, which serves as a platform for autophagosome formation<sup>10</sup>. GFP-DFCP1 also shifted to the MAM under starved conditions (Fig. 1c). These results suggest that the ATG14 complex, as well as DFCP1, relocates to the MAM fractions during starvation.

An additional marker, ATG5, was used to observe autophagosome formation. ATG5 binds only to the isolation membrane (which eventually becomes an autophagosome) and detaches from the membrane after autophagosome formation completion<sup>14</sup>; it is thus an ideal and specific marker for the initial stages of autophagosome formation. Like ATG14, we found that 52.8% of ATG5 dots were observed at the ER-mitochondria contact site (Fig. 2a and Supplementary Fig. 2b, c).

Live-cell imaging was performed with three cameras to observe three markers simultaneously. Despite the fact that ATG5 is not an ER-resident protein, 98% of ATG5 dots were found on the ER after starvation (Fig. 2c). In fact, during the entire autophagosome formation process, most ATG5 dots never separated from the ER-mitochondria contact site (Fig. 2b, Supplementary Fig. 4 and Supplementary Videos 1–3). According to the live-cell imaging analysis, 98% of ATG5 dots localized on the ER and 79% were within 0.88  $\mu$ m of the mitochondria (Fig. 2c, d). Localization near both organelles remained throughout the autophagosome formation process (Fig. 2e, f). However, unlike the stable association with the ER, association with mitochondria dynamically oscillated (but remained within 0.5  $\mu$ m), which may explain why only 50% of ATG5 dots were observed on mitochondria of fixed cells. These observations show that the ER is the platform of autophagosome formation, and dynamic mitochondria association may have a role in providing components required for this process.

When the mitochondrial protein voltage-dependent anion channel 1 (VDAC1)<sup>15</sup> was overexpressed, it localized at ER-mitochondria contact sites, indicating its viability as an ER-mitochondria contact-site marker (Supplementary Fig. 5a). In live-cell imaging, GFP-ATG5 dots often appeared next to mCherry-VDAC1 dots and subsequently disappeared with time, whereas mCherry-VDAC1 dots remained, indicating the release of ATG5 from the ER-mitochondria contact site after completion of the autophagosome formation (Supplementary Fig. 5b and Supplementary Video 4). These results strongly suggest that autophagosome formation takes place at the ER-mitochondria contact site.

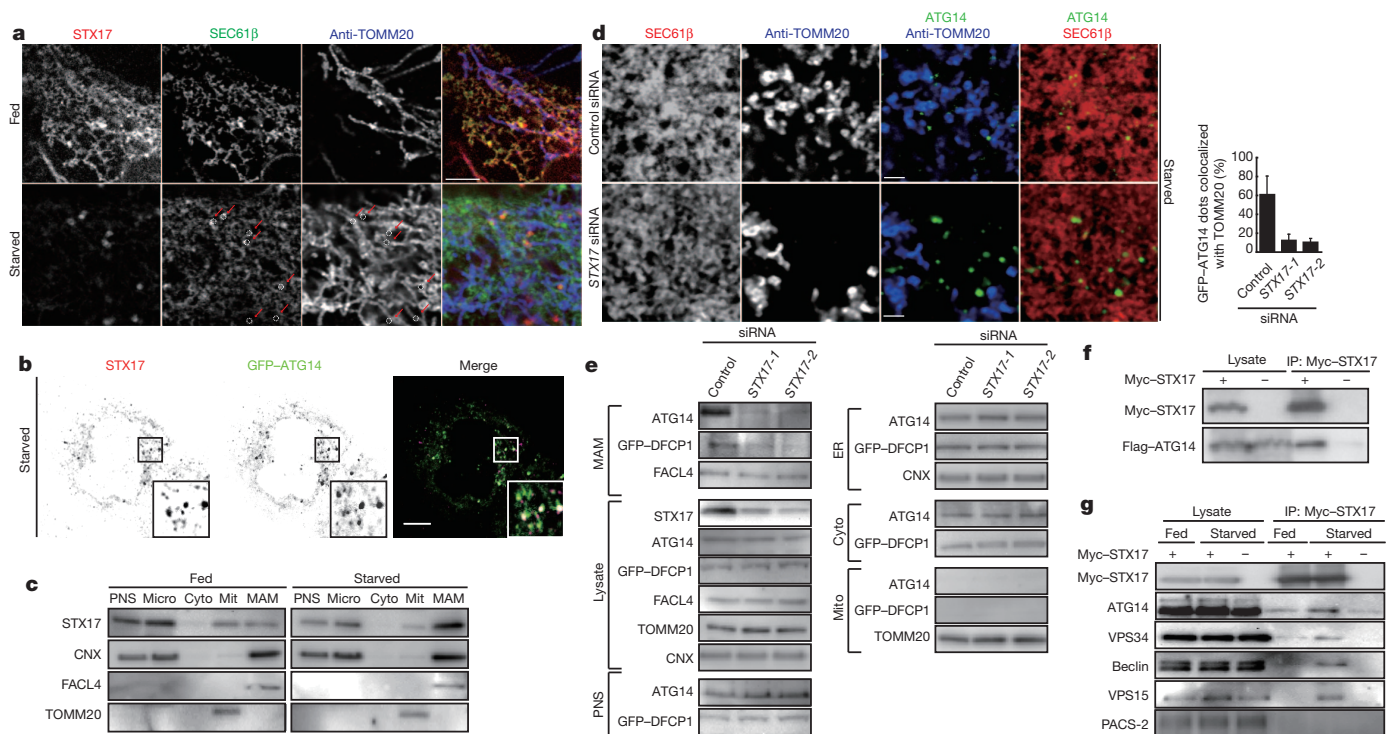
Phosphofurin acidic cluster sorting protein-2 (PACS-2), a cytosolic-sorting protein, has been implicated in the formation of ER-mitochondria contacts, and its knockdown uncouples mitochondria from the ER<sup>12</sup>. Puncta structures of GFP-ATG14 were barely visible in cells transfected with PACS2 short interfering RNA (siRNA), a similar result to the phenotype of cells deficient in mitofusin 2 (MFN2), which tethers the ER to mitochondria<sup>16,17</sup> (Supplementary Fig. 6a). In starved PACS-2 knockdown and MFN2 knockdown cells, puncta structures of GFP-LC3, which represent the isolation membrane/autophagosome/autolysosome<sup>18</sup>, were attenuated (Supplementary Fig. 6b). Indeed, the knockdown of PACS-2 caused a decrease in the lipidated form of LC3-II, which binds to autophagic membranes, and inhibited its degradation by autophagy (Supplementary Fig. 6c, d). Furthermore, siRNA knockdown of PACS-2 significantly decreased the levels of FACL4, ATG14 and GFP-DFCP1 in the MAM fraction, whereas only negligible changes were observed in other fractions (Supplementary

**Figure 2 | Isolation membrane forms at ER-mitochondria contact sites.**

**a**, COS7 cells stably expressing yellow fluorescent protein (YFP)-ATG5 were transfected with RFP-SEC61β and teal fluorescent protein (TFP)-mito, and immunostained for TOMM20. **b**, Time-lapse images of COS7 cells transfected with

YFP-ATG5, RFP-SEC61β and TFP-mito under starved conditions.

**c**, **d**, Distribution of distances of YFP-ATG5 dots from the centre to the surface of the ER (**c**) and to the surface of mitochondria (**d**). **e**, **f**, Distances of YFP-ATG5 dots over their lifetime for data from **c** (**e**) and **d** (**f**), respectively. All scale bars, 2  $\mu$ m.



**Figure 3 | STX17 regulates the relocation of ATG14 to MAM.** **a**, COS7 cells expressing GFP-SEC61 $\beta$  and STX17 were immunostained for TOMM20. **b**, HeLa cells expressing GFP-ATG14 were immunostained for STX17. **c**, Immunoblot of subcellular fractions of HEK293 cells. **d**, STX17 knockdown COS7 cells expressing GFP-ATG14 and RFP-SEC61 $\beta$  were immunostained for TOMM20. Error bars denote s.d. **e**, Immunoblot of subcellular fractions of

STX17 knockdown (using two STX17 siRNAs, STX17-1 and STX17-2) HEK293 cells expressing GFP-DFCP1. **f**, Co-immunoprecipitation of STX17 and ATG14 using HeLa cells expressing Flag-ATG14 and Myc-STX17. **g**, Pull-down of ATG14 complex from HeLa cells expressing Myc-STX17 or Myc. Scale bars, 5  $\mu$ m (**a**, **b**) and 2  $\mu$ m (**d**).

Fig. 6e). These findings were also observed in MFN2 knockdown cells (Supplementary Fig. 6f), and suggest that the inhibition of the relocation of ATG14 and DFCP1 to the ER-mitochondria contact site prevents the initiation of autophagosome formation.

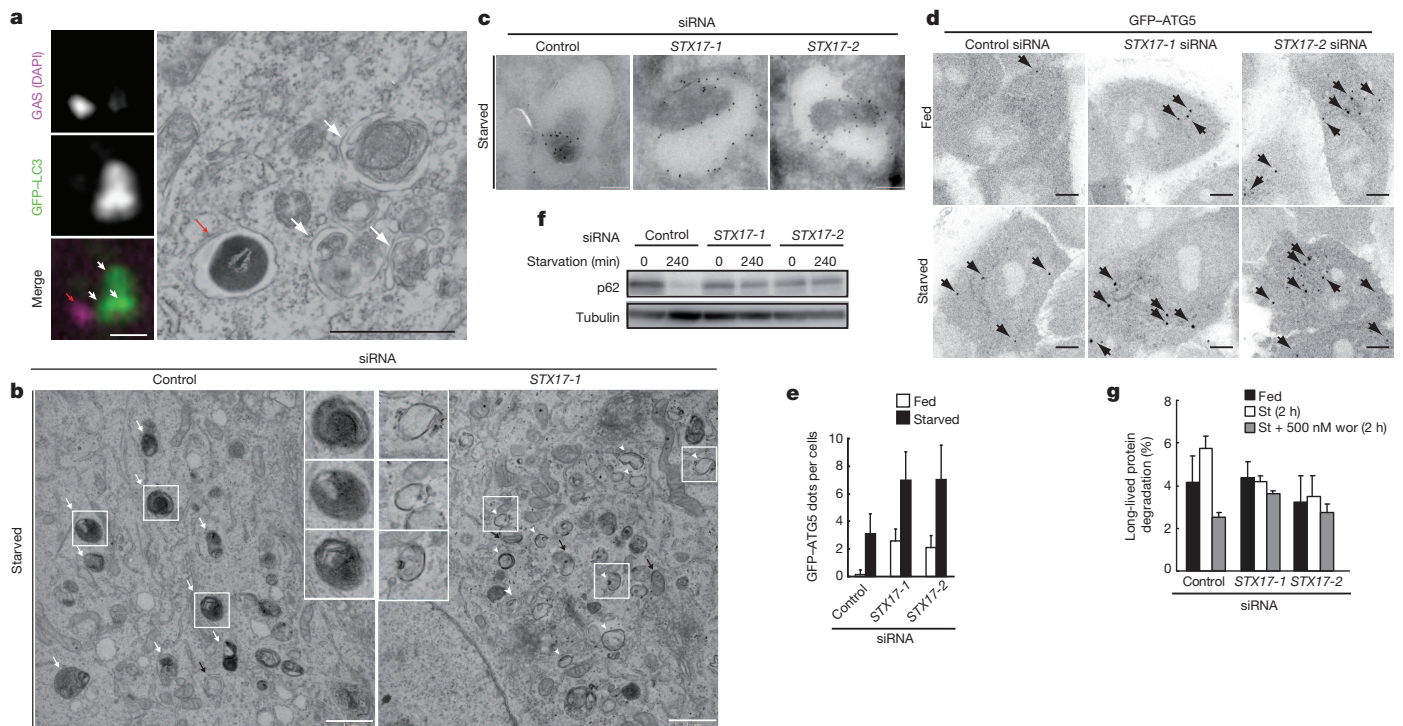
Along with ATG14 and DFCP1, another ER protein, STX17, was seen to relocate to the ER-mitochondria contact sites after starvation (Fig. 3a). We identified STX17 in a screen for ER-resident SNARE proteins involved in autophagy against invading group A *Streptococcus* (GAS)<sup>19,20</sup> (Supplementary Fig. 7). STX17 is known to be a Qa-SNARE protein (a subfamily of Q (glutamate-containing)-SNARE proteins that occupy the same position as STX1A in the SNARE complex), but its function is not known<sup>21</sup>. STX17 dots at ER-mitochondria contact sites were extensively colocalized with ATG14 under starvation conditions (Fig. 3b). Furthermore, STX17 was abundant in microsome fractions under fed conditions, whereas it was mostly enriched in the MAM fraction of starved cells (Fig. 3c). Neither PACS-2 knockdown nor MFN2 knockdown had an effect on STX17 expression, but depletion of both inhibited the relocation of STX17 to the MAM fraction (Supplementary Fig. 6e, f).

GFP-ATG14 dots did not locate at the ER-mitochondria contact site in STX17 knockdown cells under starved conditions, although they did colocalize with other part of the ER (Fig. 3d). Consistently, ATG14 and GFP-DFCP1 were clearly decreased in the MAM fraction (Fig. 3e). By contrast, ATG14 depletion exhibited no effect on starvation-induced relocation of STX17 to the ER-mitochondria contact site (Supplementary Fig. 8). These data indicate that STX17 acts upstream of ATG14, and is responsible for directing ATG14 to the ER-mitochondria contact site. Given the functional relationship between STX17 and ATG14, we performed co-immunoprecipitation experiments. Myc-STX17 did not significantly co-immunoprecipitate with GFP-DFCP1 under starved conditions, whereas it did pull down Flag-ATG14 (Fig. 3f and data not shown). Similarly, endogenous ATG14 as well as other components of the complex clearly

co-immunoprecipitated with Myc-STX17 under starved conditions, whereas ATG14 barely precipitated in the absence of Myc-STX17 (Fig. 3g). These results suggest that Myc-STX17 associates with ATG14 only after starvation.

Whereas STX17 knockdown hinders the recruitment of ATG14 to the ER-mitochondria contact site, ATG14 was still able to assemble at other locations (Fig. 3d), indicating that STX17 is necessary for the recruitment of ATG14 to the ER-mitochondria contact site but not for the assembly of ATG14 dots. LC3-positive puncta were formed and accumulated in STX17 knockdown cells, indicating that the autophagosome maturation step was arrested (Supplementary Fig. 10b). We examined LC3 puncta by correlative light and electron microscopy in STX17 knockdown cells infected by GAS, and found that LC3-positive puncta represent isolation membranes (Fig. 4a and Supplementary Fig. 9). Immunogold electron microscopy of GFP-LC3, as well as conventional electron microscopy, showed accumulation of isolation membranes in STX17 knockdown cell under starvation conditions (Fig. 4b, c). In control cells, autolysosomes (autophagosomes fused with lysosomes) but few isolation membranes were observed; by contrast, in STX17 knockdown cells, isolation membranes were dominant whereas autolysosomes were nearly absent (Fig. 4c). Together, these data suggested that STX17 knockdown leads to an accumulation of isolation membranes deficient in autophagosome maturation. In addition, GFP-ATG5 dots accumulated to a considerable degree in these cells even under fed conditions, similar to the phenotype of inhibited autophagosome completion in ATG4B mutants<sup>22</sup> (Fig. 4d, e). Consistent with this, knockdown of STX17 significantly decreased autophagy, as detected by several assays: (1) degradation of p62 (also known as SQSTM1)<sup>23</sup> (Fig. 4f); (2) degradation of long-lived proteins<sup>22</sup>, the clearance of which depends on autophagy (Fig. 4g); (3) autophagic flux as measured by the amount of LC3 in the presence or absence of lysosomal protease inhibitors<sup>24</sup> (Supplementary Fig. 10a); (4) fusion between autophagosomes and lysosomes, as measured by





**Figure 4 | Functional autophagosomes do not form in STX17 knockdown cells.** **a**, Correlative light and electron microscopy analysis of STX17 knockdown HeLa cells expressing GFP-LC3 (white arrows) infected with GAS (red arrows). DAPI, 4',6-diamidino-2-phenylindole. **b**, Electron micrographs of STX17 knockdown cells. White arrows denote autolysosomes, black arrows denote autophagosome, white arrowheads mark isolation membranes.

monomeric red fluorescent protein (mRFP)–GFP tandem-tagged LC3 (ref. 25) (Supplementary Fig. 10b); (5) formation of large autophagosomes engulfing GAS (Supplementary Fig. 9); and (6) autophagic killing of GAS (Supplementary Fig. 7c). These results indicate that knockdown of STX17 blocks the completion of the autophagosome formation. Therefore, it is likely that STX17-dependent recruitment of ATG14 to the MAM is essential for the formation of functional autophagosomes. Our most plausible explanation for the phenotypic differences between knockdown of STX17 and knockdown of PACS-2 or MFN2 is the existence of STX17. In PACS-2 or MFN2 knockdown cells, the disruption of the MAM allows STX17 to interact with ATG14 (Supplementary Fig. 6g), blocking ATG14 dot formation. In STX17 knockdown cells, the absence of STX17 frees ATG14 to initiate isolation membrane formation.

Our findings provide new insights into two competing models for autophagosome biogenesis, that is, origin in the ER<sup>7–9</sup> or the mitochondria<sup>16</sup>. Results that have exclusively supported either model can now be explained by a new model in which the autophagosome is formed at the ER–mitochondria contact site (Supplementary Fig. 1). In retrospect, it seems that research groups arguing in favour of either single organelle model have been observing the same phenomenon from different sides. Our model does not exclude a possibility that formation independent on the contact site could still occur. We have also discovered a novel role for the ER–mitochondria contact site, a special crossing-over point between two organelles that is involved in a range of physiological functions. Many questions still remain, but these findings open new avenues for further study of the molecular mechanisms of autophagy.

## METHODS SUMMARY

All cell lines were cultured in DMEM (Wako) supplemented with 10% heat-inactivated FBS (Invitrogen). siRNA and siRNA negative control (Stealth RNAi) were purchased from Invitrogen. Cells were examined under a fluorescence laser

scanning confocal microscope (LSM510; Carl Zeiss and FV1000; Olympus) using the Zeiss LSM Image Browser software (Carl Zeiss) and FLUOVIEW (Olympus). Live-cell imaging was performed with the Yokogawa spinning disc confocal system (CSU-X1; Olympus). Immunoblotting<sup>22</sup>, isolation of the MAM<sup>3</sup>, conventional electron microscopy<sup>22</sup>, immunoelectron microscopy<sup>26</sup>, and the bulk protein degradation assay<sup>22</sup> were performed as described previously. All data are presented as mean  $\pm$  s.d.

**Full Methods** and any associated references are available in the online version of the paper.

Received 26 September 2011; accepted 15 January 2013.

Published online 3 March 2013.

- Mizushima, N., Yoshimori, T. & Ohsumi, Y. The role of Atg proteins in autophagosome formation. *Annu. Rev. Cell Dev. Biol.* **27**, 107–132 (2011).
- Tooze, S. A. & Yoshimori, T. The origin of the autophagosomal membrane. *Nature Cell Biol.* **12**, 831–835 (2010).
- Vance, J. E. Phospholipid synthesis in a membrane fraction associated with mitochondria. *J. Biol. Chem.* **265**, 7248–7256 (1990).
- Rizzuto, R. *et al.* Close contacts with the endoplasmic reticulum as determinants of mitochondrial  $\text{Ca}^{2+}$  responses. *Science* **280**, 1763–1766 (1998).
- Hayashi, T., Rizzuto, R., Hajnoczky, G. & Su, T.-P. MAM: more than just a housekeeper. *Trends Cell Biol.* **19**, 81–88 (2009).
- Matsunaga, K. *et al.* Two Beclin 1-binding proteins, Atg14L and Rubicon, reciprocally regulate autophagy at different stages. *Nature Cell Biol.* **11**, 385–396 (2009).
- Matsunaga, K. *et al.* Autophagy requires endoplasmic reticulum targeting of the PI3-kinase complex via Atg14L. *J. Cell Biol.* **190**, 511–521 (2010).
- Hayashi-Nishino, M. *et al.* A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation. *Nature Cell Biol.* **11**, 1433–1437 (2009).
- Ylä-Anttila, P., Vihinen, H., Jokitalo, E. & Eskelinen, E.-L. 3D tomography reveals connections between the phagophore and endoplasmic reticulum. *Autophagy* **5**, 1180–1185 (2009).
- Axe, E. L. *et al.* Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. *J. Cell Biol.* **182**, 685–701 (2008).
- Hayashi, T. & Su, T.-P. Sigma-1 receptor chaperones at the ER-mitochondrion interface regulate  $\text{Ca}^{2+}$  signaling and cell survival. *Cell* **131**, 596–610 (2007).
- Simmen, T. *et al.* PACS-2 controls endoplasmic reticulum-mitochondria communication and Bid-mediated apoptosis. *EMBO J.* **24**, 717–729 (2005).
- Künkele, K. P. *et al.* The preprotein translocation channel of the outer membrane of mitochondria. *Cell* **93**, 1009–1019 (1998).



14. Mizushima, N. *et al.* Dissection of autophagosome formation using Apg5-deficient mouse embryonic stem cells. *J. Cell Biol.* **152**, 657–668 (2001).
15. Szabadkai, G. *et al.* Chaperone-mediated coupling of endoplasmic reticulum and mitochondrial  $\text{Ca}^{2+}$  channels. *J. Cell Biol.* **175**, 901–911 (2006).
16. Hailey, D. W. *et al.* Mitochondria supply membranes for autophagosome biogenesis during starvation. *Cell* **141**, 656–667 (2010).
17. de Brito, O. M. & Scorrano, L. Mitofusin 2 tethers endoplasmic reticulum to mitochondria. *Nature* **456**, 605–610 (2008).
18. Kabeya, Y. *et al.* LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *EMBO J.* **19**, 5720–5728 (2000).
19. Furuta, N., Fujita, N., Noda, T., Yoshimori, T. & Amano, A. Combinational soluble N-ethylmaleimide-sensitive factor attachment protein receptor proteins VAMP8 and Vti1b mediate fusion of antimicrobial and canonical autophagosomes with lysosomes. *Mol. Biol. Cell* **21**, 1001–1010 (2010).
20. Steegmaier, M., Oorschot, V., Klumperman, J. & Scheller, R. H. Syntaxin 17 is abundant in steroidogenic cells and implicated in smooth endoplasmic reticulum membrane dynamics. *Mol. Biol. Cell* **11**, 2719–2731 (2000).
21. Fujita, N. *et al.* An Atg4B mutant hampers the lipidation of LC3 paralogues and causes defects in autophagosome closure. *Mol. Biol. Cell* **19**, 4651–4659 (2008).
22. Bjørkøy, G. *et al.* Monitoring autophagic degradation of p62/SQSTM1. *Methods Enzymol.* **452**, 181–197 (2009).
23. Mizushima, N. & Yoshimori, T. How to interpret LC3 immunoblotting. *Autophagy* **3**, 542–545 (2007).
24. Kimura, S., Noda, T. & Yoshimori, T. Dissection of the autophagosome maturation process by a novel reporter protein, tandem fluorescent-tagged LC3. *Autophagy* **3**, 452–460 (2007).
25. Yamamoto, A. & Masaki, R. Pre-embedding nanogold silver and gold intensification. *Methods Mol. Biol.* **657**, 225–235 (2010).
26. Kageyama, S. *et al.* The LC3 recruitment mechanism is separate from Atg9L1-dependent membrane formation in the autophagic response against Salmonella. *Mol. Biol. Cell* **22**, 2290–2300 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank all members of the Amano and Yoshimori laboratories for discussions. We thank O. Kunitaki and K. Miura for continuous help throughout the research. We also thank P. Karagiannis for proofreading. This research was supported by grants-in-aid for Scientific Research (A) and (B) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

**Author Contributions** M.H., N. Furuta, T.H., Y.H., N. Fujita, T.N., T.Y. and A.A. designed the experiments. M.H. performed image analysis, including confocal and time-lapse microscopic analysis. A.M. performed image analysis on live-cell imaging. A.Y., M.H. and A.N. performed immunoelectron microscopy. M.H. and A.N. performed conventional electron microscopy. H.O. performed correlative light and electron microscopy. N. Furuta performed the remaining experiments. M.H., N. Furuta, T.Y. and A.A. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.Y. ([tamyoshi@fbs.osaka-u.ac.jp](mailto:tamyoshi@fbs.osaka-u.ac.jp)) or A.A. ([amanoa@dent.osaka-u.ac.jp](mailto:amanoa@dent.osaka-u.ac.jp)).

## METHODS

**Cell culture and transfection.** All cell lines were cultured in DMEM (Wako) supplemented with 10% heat-inactivated FBS (Invitrogen). HeLa cells stably expressing enhanced GFP (eGFP)-LC3 and COS7 cells stably expressing YFP-ATG5 were constructed as previously described<sup>14</sup>. HeLa cells stably expressing eGFP-ATG5 were a gift from N. Fujita. Wild-type and ATG14-deficient mouse embryonic fibroblasts were constructed as previously described<sup>26</sup>. HEK293 cells stably expressing eGFP-DFCP1 were established as previously described<sup>10</sup>. LipofectAMINE2000 (Invitrogen) was used for transfection. For amino acid starvation, cells were cultured in EBSS (Sigma) without amino acids or FBS for 2 h. Stable transformants were selected in complete medium containing 500 µg ml<sup>-1</sup> G418 (Sigma).

**Antibodies and immunoblotting.** Antibodies against the following proteins were mouse monoclonal: Myc (MBL), GFP (Abcam), tubulin (Sigma), PDI (Stressgen) and CNX (BD Bioscience). Rabbit polyclonals include: LC3 (MBL), TOMM20 (Abcam), FACL4 (Abcam), STX17 (Sigma), ATG14 (MBL), VPS34 (Cell Signaling), VPS15 (MBL), beclin (MBL), MFN2 (Sigma) and PACS-2 (ProteinTech Group). Alexa Fluor-conjugated secondary antibodies (goat anti-mouse IgG and goat anti-Rabbit IgG) used for fluorescence microscopy were purchased from Invitrogen. Immunoblotting was performed as previously described<sup>27</sup>.

**siRNAs and plasmids.** siRNA duplexes targeting *STX17* (HSS123730 and HSS123732), *STX18* (HSS122692 and HSS122693), *SEC20* (also known as *BNIP1*) (HSS141385 and HSS141386), *SEC22A* (HSS120253 and HSS120254), *SEC22C* (HSS145024 and HSS145025), *SLT1* (also known as *USE1*) (HSS125167 and HSS125168), *PACS2* (HSS146278 and HSS146279), *ATG14* (HSS117763 and HSS117764), *MFN2* (HSS115028 and HSS115029), and siRNA negative control (Stealth RNAi) were purchased from Invitrogen. siRNA for *VTI1B* was described previously<sup>19</sup>.

Expression vectors for mCherry-C1 plasmids<sup>5</sup> and mRFP-GFP-LC3 (ref. 24) plasmids were constructed. GFP-ATG14 and Flag-ATG14 were also established<sup>7</sup>. mCherry- or GFP-Cyt b5-TM was constructed as described previously<sup>28</sup>. GFP-SEC61β was a gift from T. Rapoport. DsRed-ER was a gift from K. Tabata. DsRed2-mito was obtained from Clontech. TFP-mito was constructed by replacing DsRed using PCR fragments. To construct Myc-STX17 plasmids, complementary DNA cloned from genomic DNA of HeLa cells was inserted into pcDNA3.1/Myc-His(-)A (Invitrogen) using engineered EcoRI and XhoI sites. To construct Myc-STX17(ΔSNARE) plasmids, the STX17(ΔSNARE) domain was generated by PCR (primers: 5'-ATGCTCTGAAGATGAAGAAAAAGTG-3' and 5'-TTAAGCCTTCCCTAAGTTTGTGG-3'). The PCR fragment was digested with EcoRI and XhoI and used to replace the *STX17* cDNA in the Myc-STX17 plasmid. To construct Myc-ATG14 plasmids, cDNA from Flag-ATG14 was inserted into pcDNA3.1/Myc-His(-)A (Invitrogen) using engineered EcoRI and KpnI sites. To construct mCherry or GFP-VDAC1 plasmids, cDNA from genomic DNA of HeLa cells was inserted into mCherry-C1 or EGFP-C1 (Invitrogen) using engineered SalI and BamHI sites. To construct mCherry or GFP-DRP1 plasmids, cDNA from genomic DNA of HeLa cells was inserted into mCherry-C1 or EGFP-C1 (Invitrogen) using engineered SalI and BamHI sites. To construct an mCherry-LC3 plasmid, PCR was performed to generate mCherry cDNA with exogenous restriction sites for BamHI and NotI at the 5' and 3' ends, respectively. The PCR fragment obtained after digestion with each restriction enzyme was inserted into pMRX-mCherry. LC3 cDNA cloned from the genomic DNA of HeLa cells was inserted into pMRX-mCherry using engineered EcoRI and XhoI sites.

**RT-PCR and real-time PCR.** PCR with reverse transcription (RT-PCR) and real-time PCR were performed as described previously<sup>29</sup>. The sequences of the oligonucleotides were as follows: *STX17* (forward, 5'-GCCGTCTTGAACAGCT ATC-3'; reverse, 5'-TTCATGCAACTGTCCAG-3'), *STX18* (forward, 5'-TCACATTGGCAAAGTGAAGAT-3'; reverse, 5'-TCTGGGCATCTGGTC TATC-3'), *SEC20* (forward, 5'-ACGTCCGGATCTGTAACCAA-3'; reverse, 5'-AAGGGTCTGAACAATCAGC-3'), *SEC22A* (forward, 5'-ATGCTTTCGAG GAACTTGC-3'; reverse, 5'-GAGAAGGCGAGAACATTTGG-3'), *SEC22C* (forward, 5'-GAGTTTAGCCTTGGGACTGG-3'; reverse, 5'-AGGAGCAGAT AGCCATGCA-3'), *SLT1* (forward, 5'-GTCCAGGCTGGAGCTAAACC-3'; reverse, 5'-GGCCTGCAACATGTCTCTA-3'), and *GAPDH* (forward, 5'-TGTTCTCTGACTTCAACAG-3'; reverse, 5'-CTGTAGCCAAATTCGT TGTCATAC-3'). The expression levels of messenger RNA are indicated as the relative cycle number for a given gene normalized against the cycle number for *GAPDH*. Each reaction was repeated at least three times to assess reproducibility.

**Fluorescence microscopy.** Fluorescence microscopy was performed as described previously<sup>5</sup>. To detect STX17, cells were permeabilized with 0.1% Triton X-100, and blocked with 1% normal goat serum (Abcam).

**Proteinase inhibitor.** E64d and pepstatin A were purchased from the Peptide Institute. Cells were transfected with siRNA against *STX17* or with control siRNA. After 48 h of incubation, the cells were cultured in EBSS containing E64d

(10 µg ml<sup>-1</sup>) and pepstatin A (30 µg ml<sup>-1</sup>) or in a solution lacking inhibitors (control; dimethylsulphoxide 1.1 µg ml<sup>-1</sup>) at the appropriate times. Lysates were examined by western blotting using antibodies against LC3 and tubulin.

**GAS infection and bacterial viability assay.** Infection with GAS (strain JRS4) and the colony-forming units (c.f.u.) viability assay were performed as described previously<sup>30</sup>.

**Immunoprecipitation analysis.** Immunoprecipitation analysis was performed using the c-Myc-tagged protein mild purification kit (MBL). Twenty-million cells were suspended in lysis buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA and 1% CHAPS). Cell lysates were centrifuged at 15,000g for 10 min. The resulting supernatant was transferred to a spin column and incubated with anti-c-Myc beads for 1 h at 4 °C. Anti-c-Myc beads in the spin column were incubated with elution peptide solution for 30 s on ice and centrifuged for 10 s.

**Electron microscopy and bulk protein degradation assay.** Conventional electron microscopy was performed as described previously<sup>31</sup>. Immunoelectron microscopy was performed using a rabbit polyclonal antibody against GFP (a gift from N. Nakamura) by applying the pre-embedding gold enhancement method as described previously<sup>25</sup>. Samples were analysed with an electron microscope (H7600; Hitachi). Correlative light and electron microscopy was performed as described previously<sup>26</sup>, as was the long-lived proteins degradation assay<sup>21</sup>. In brief, degradation of long-lived proteins depends on the autophagic activity. Cells were seeded in 24-well dishes and incubated overnight. On the next day, the cells were exchanged into labelling medium containing <sup>14</sup>C-valine (1.5 Ci ml<sup>-1</sup>) and again incubated overnight. Cells were exchanged into chase medium (DMEM supplemented with 10% FBS and 10 mM unlabelled valine) and further incubated for 4 h to remove the contribution of short-lived proteins. After the chase period, media was replaced with either growth medium containing 10 mM valine or EBSS containing 10 mM valine (to induce autophagy). After 2 h incubation, the media were collected, and the trichloroacetic acid (TCA)-soluble fraction was analysed by scintillation counting. The cells were lysed in ice-cold RIPA buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 1% deoxycholate, 5 mM EDTA and protease inhibitor cocktail (Roche)); the TCA-insoluble fraction was isolated and analysed by scintillation counting. To determine the rate of long-lived protein degradation, the count in the TCA-soluble fraction in the medium was divided by the total cellular count.

**Tokuyasu sampling and immunolabelling.** Cells transfected with Myc-ATG14 were pre-fixed for 10 min (4% paraformaldehyde/0.1% glutaraldehyde in PHEM (120 mM PIPES, 50 mM HEPES, 20 mM EGTA and 4 mM MgCl<sub>2</sub>, pH 6.9)) then fixed with fresh 4% paraformaldehyde in PHEM for 30 min at room temperature. Cells were scraped, resuspended in 10% gelatin/PBS, pelleted by centrifugation, and immediately placed on ice until the gelatin solidified. Gelatin-embedded samples were infiltrated with 2.3 M sucrose (in 0.1 M phosphate buffer), placed at 4 °C overnight on a rotating wheel mounted onto sample pins, and frozen in liquid nitrogen. Subsequently, samples were cryo-sectioned at 60 nm thickness using an FC6/UC6 cryo-ultramicrotome (Leica), and selected with a 1:1 mixture of 2% methylcellulose (0.025 Pa s<sup>-1</sup>; Sigma) and 2.3 M sucrose. After thawing, sections were first reacted with anti-Myc antibodies (MBL) diluted 1:100, then with rabbit anti-mouse antibodies (Invitrogen) diluted 1:1,000, and finally with 10-nm gold-conjugated protein A following an established protocol<sup>32</sup>. For double labelling, an anti-PDI antibody was used (Stressgen). Sections were embedded in 4% uranyl acetate/2% methylcellulose (ratio 1:9)<sup>33</sup> and viewed using a JEOL 1010 electron microscope.

**Live-cell imaging.** Live-cell imaging was performed at 37 °C using an Olympus IX81 microscope with a ×100 immersion objective lens (numerical aperture 1.40) equipped with a Yokogawa spinning disc confocal system (CSU-X1) and three EMCCD cameras (Andor iXon3). Images were acquired by Andor iQ live cell imaging software.

**Image analysis.** Images were digitally processed as follows for subsequent statistical analyses. Bleed-through of the CFP channel in the YFP channel was removed by spectral unmixing. The background fluorescence in the YFP channel was then normalized using time-averaged images. Next, all channels were processed by a denoising algorithm to improve the signal-to-noise ratio in the fluorescence images<sup>34</sup> and finally by constrained iterative deconvolution. The RFP channel is shown in logarithmic scale to visualize both the peripheral mesh structures and the thick network around the nucleus. The three-dimensional distance between the centre of a YFP-ATG5 dot and the surface of the ER/mitochondria was measured with custom-made software. The mitochondria and ER had relatively constant fluorescence intensities and thus adaptive thresholding was effective to distinguish their surfaces. On the other hand, because YFP-ATG5 dots have variable signal intensities, thresholding alone can change their size. Therefore, to reduce thresholding artefacts, centres rather than surface coordinates of the YFP-ATG5 dots were used. A surface of the ER/mitochondria complex was represented as the nearest pixel to the centre of a YFP-ATG5 dot.

27. Kamimoto, T. *et al.* Intracellular inclusions containing mutant  $\alpha$ 1-antitrypsin Z are propagated in the absence of autophagic activity. *J. Biol. Chem.* **281**, 4467–4476 (2006).
28. Saitoh, T. *et al.* Atg9a controls dsDNA-driven dynamic translocation of STING and the innate immune response. *Proc. Natl Acad. Sci. USA* **106**, 20842–20846 (2009).
29. Kawai, S., Yamauchi, M., Wakisaka, S., Ooshima, T. & Amano, A. Zinc-finger transcription factor odd-skipped related 2 is one of the regulators in osteoblast proliferation and bone formation. *J. Bone Miner. Res.* **22**, 1362–1372 (2007).
30. Nakagawa, I. *et al.* Autophagy defends cells against invading group A *Streptococcus*. *Science* **306**, 1037–1040 (2004).
31. Fujita, N. *et al.* The Atg16L complex specifies the site of LC3 lipidation for membrane biogenesis in autophagy. *Mol. Biol. Cell* **19**, 2092–2100 (2008).
32. Slot, J. W. & Geuze, H. J. Cryosectioning and immunolabeling. *Nature Protocols* **2**, 2480–2491 (2007).
33. Griffiths, G., McDowall, A., Back, R. & Dubochet, J. On the preparation of cryosections for immunocytochemistry. *J. Ultrastruct. Res.* **89**, 65–78 (1984).
34. Boulanger, J., Kervrann, C. & Bouthemy, P. A simulation and estimation framework for intracellular dynamics and trafficking in video-microscopy and fluorescence imagery. *Med. Image Anal.* **13**, 132–142 (2009).



# Multidomain integration in the structure of the HNF-4 $\alpha$ nuclear receptor complex

Vikas Chandra<sup>1</sup>, Pengxiang Huang<sup>1</sup>, Nalini Potluri<sup>1</sup>, Dalei Wu<sup>1</sup>, Youngchang Kim<sup>2</sup> & Fraydoon Rastinejad<sup>1</sup>

The hepatocyte nuclear factor 4 $\alpha$  (HNF-4 $\alpha$ ; also known as NR2A1) is a member of the nuclear receptor (NR) family of transcription factors, which have conserved DNA-binding domains and ligand-binding domains<sup>1,2</sup>. HNF-4 $\alpha$  is the most abundant DNA-binding protein in the liver, where some 40% of the actively transcribed genes have a HNF-4 $\alpha$  response element<sup>1,3,4</sup>. These regulated genes are largely involved in the hepatic gluconeogenic program and lipid metabolism<sup>3,5,6</sup>. In the pancreas HNF-4 $\alpha$  is also a master regulator, controlling an estimated 11% of islet genes<sup>7</sup>. HNF-4 $\alpha$  protein mutations are linked to maturity-onset diabetes of the young, type 1 (MODY1) and hyperinsulinaemic hypoglycaemia<sup>8–11</sup>. Previous structural analyses of NRs, although productive in elucidating the structure of individual domains, have lagged behind in revealing the connectivity patterns of NR domains. Here we describe the 2.9 Å crystal structure of the multidomain human HNF-4 $\alpha$  homodimer bound to its DNA response element and coactivator-derived peptides. A convergence zone connects multiple receptor domains in an asymmetric fashion, joining distinct elements from each monomer. An arginine target of PRMT1 methylation protrudes directly into this convergence zone and sustains its integrity. A serine target of protein kinase C is also responsible for maintaining domain–domain interactions. These post-translational modifications lead to changes in DNA binding by communicating through the tightly connected surfaces of the quaternary fold. We find that some MODY1 mutations, positioned on the ligand-binding domain and hinge regions of the receptor, compromise DNA binding at a distance by communicating through the interjunctional surfaces of the complex. The overall domain representation of the HNF-4 $\alpha$  homodimer is different from that of the PPAR- $\gamma$ -RXR- $\alpha$  heterodimer, even when both NR complexes are assembled on the same DNA element. Our findings suggest that unique quaternary folds and inter-domain connections in NRs could be exploited by small-molecule allosteric modulators that affect distal functions in these polypeptides.

We previously reported the only high-resolution structural example of a multidomain NR complex, that of the PPAR- $\gamma$ -RXR- $\alpha$  heterodimer on its DNA response element<sup>12</sup>. To understand the extent of domain integration in other NRs, here we analyse the crystal structure of the complex of HNF-4 $\alpha$ , an obligate homodimer, bound to its DNA element and coactivator-derived peptides. HNF-4 $\alpha$  uses the linear domain arrangement shown in Fig. 1a. Our efforts to crystallize full-length HNF-4 $\alpha$  were unsuccessful. However, by proteolytically probing its DNA-assembled complex, we identified an extended segment comprising the DNA-binding domain (DBD)–hinge–ligand-binding domain (LBD) portions and corresponding to residues 46–368 (Supplementary Fig. 1). Cloning, expression and purification of the stable DBD–hinge–LBD multidomain segment made it possible to obtain well-diffracting crystals of a complex with its consensus response element and coactivator (NCOA2) peptide. Electron density maps for all the critical junctions of the complex are shown in Supplementary Figs 2–5. The response element consists of a direct repeat of AGGTCA half-sites with one base-pair spacing (DR1). The

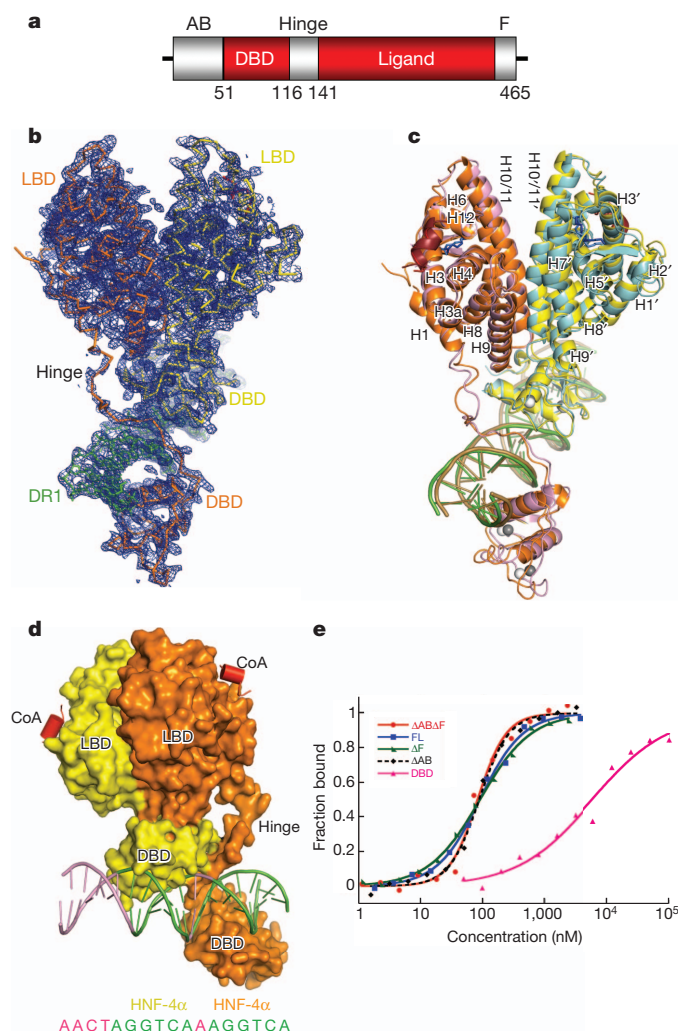
DR1 is the major consensus binding site for both HNF-4 $\alpha$  and PPAR- $\gamma$ -RXR- $\alpha$ <sup>4,6,13</sup>.

X-ray diffraction data was collected to 2.9 Å resolution, and the structure refined (see Supplementary Table 1). The crystal asymmetric unit contains two independent representations of the HNF-4 $\alpha$ -homodimer–DNA–peptide complex. The electron density map from one complex and the comparison of the two complexes is shown in Fig. 1b, c. The two representations are nearly identical, with root mean squared deviations of less than 2.0 Å over all their atoms. The LBD and DBD portions match their previously determined isolated structures (Supplementary Figs 6 and 7). Both DBDs are in register with their half-sites, interacting with the major grooves (Supplementary Figs 8 and 9). Helix 12 of the LBDs is in the active conformation, and a coactivator LXXLL peptide is bound to each LBD.

The HNF-4 $\alpha$  homodimer shows a striking and complex pattern of interfacial junctions. A central zone incorporates surfaces from both LBDs, the DBD of the upstream subunit, and the hinge region of the downstream subunit. This domain convergence zone suggests a path of communication between the conserved domains through their coupled surfaces (Fig. 2a). The LBDs, symmetrical in their mutual interactions when viewed in isolation, cooperate in a highly asymmetric fashion to straddle the surface of only the upstream DBD (Fig. 1d). As a result, the overall complex appears partitioned towards the upstream half of the DR1, and adopts a highly asymmetrical organization for a homodimeric transcription factor. A previous study suggested that HNF-4 $\alpha$  homodimers could bind asymmetrically to their DNA response elements<sup>6</sup>. The resulting quaternary arrangement creates precisely the correct DBD-to-DBD distances needed to match the geometric constraints of the two AGGTCA half-sites and their intervening spacer. At the same time, the quaternary organization renders both LBD pockets and their coactivator-interacting surfaces unencumbered, allowing free access to both ligands and LXXLL elements, respectively.

The interface that forms between the DBD of the upstream subunit and the hinge region of the downstream subunit is one important domain–domain interface of the complex, and is reminiscent of an interaction we described previously in the PPAR- $\gamma$ -RXR- $\alpha$  complex<sup>12</sup>. The resulting arrangement places the two DBDs in a solid head-to-tail arrangement that extends their combined footprint to match their DR1 contact surface perfectly. The manner in which the two LBDs cooperate to interact with the upstream DBD suggests that the physical integration of all three domains may be required for high-affinity DNA binding (Fig. 1d). To test this idea, we measured the DNA affinity of the HNF-4 $\alpha$  protein when various segments of the protein were removed (Fig. 1d). Measuring first the DNA-binding affinity of the HNF-4 $\alpha$  that only contains its DBD and hinge portions, we observed very weak binding to DR1 with an equilibrium dissociation constant ( $K_d$ ) of approximately 6,000 nM. When the LBD portion of the receptor is contained within the polypeptide, the complex showed a 75-fold enhanced affinity for DR1, with a  $K_d$  of approximately 80 nM (see Fig. 1e and Supplementary Table 2). These results are consistent with

<sup>1</sup>Metabolic Signaling and Disease Program, Sanford-Burnham Medical Research Institute, Orlando, Florida 32827, USA. <sup>2</sup>Biosciences Division, Structural Biology Center, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, USA.

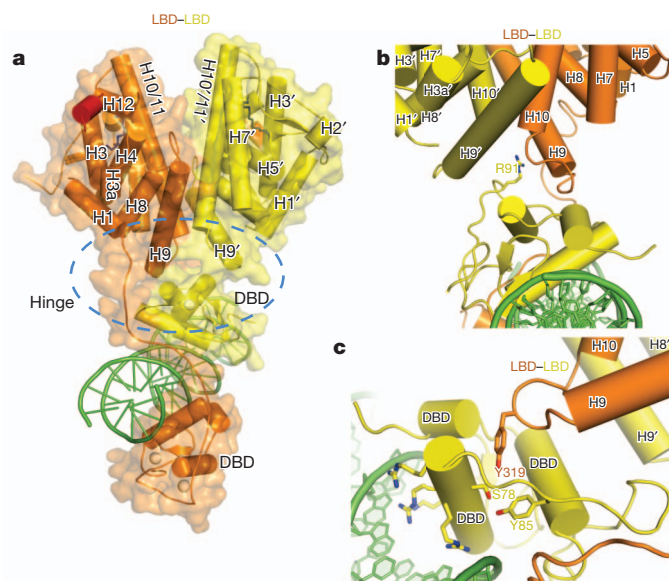


**Figure 1 | Overall organization of the HNF-4 $\alpha$  homodimer on DNA.**

**a**, Linear depiction of the HNF-4 $\alpha$  protein domains. **b**, Electron density map ( $2F_o - F_c$ ) for one of the two HNF-4 $\alpha$ -homodimer-DNA complexes of the crystallographic asymmetric unit. **c**, Superposition of the two independent HNF-4 $\alpha$ -homodimer-DNA complexes in the asymmetric unit. The two homodimeric complexes are coloured orange/yellow in one case, and blue/purple in the other. The ligands in the LBD are shown in blue, Zn is shown in grey, and the coactivator peptides are shown in red/brown. Numbers with prime refer to the upstream-positioned subunit. **d**, The positioning of both LBDs in a complex on top of the upstream DBD enables high-affinity DR1 DNA binding. **e**, Contribution of various receptor domains to the DNA-binding affinity of HNF-4 $\alpha$ . DNA binding is measured using fluorescent polarization studies with a 5' fluorescein isothiocyanate (FITC)-labelled DR1. The x-axis shows concentration of the HNF-4 $\alpha$  protein, and the y-axis shows the fractional DNA bound. Removal of the AB domain ( $\Delta AB$ ), F domain ( $\Delta F$ ) or both ( $\Delta AB\Delta F$ ) does not alter the DNA affinity compared to the full-length (FL) receptor. However, removal of the LBD reduces the affinity of the resulting DBD-hinge (DBD) region to a  $K_d$  of approximately 6,000 nM, whereas the presence of the LBD together with the DBD hinge ( $\Delta AB\Delta F$ ) allows DNA binding with a  $K_d$  of 82 nM (see also Supplementary Table 2).

our observation that the LBD and DBD modules are physically and functionally integrated to establish high-affinity DNA binding. They are also consistent with a previous study that showed that the LBD significantly enhances the half-life of the HNF-4 $\alpha$  DNA-binding complex<sup>14</sup>.

We next measured the DNA-binding contributions of the amino (N)-terminal (AB region) and carboxy (C)-terminal (F region) portions of the polypeptide, both of which were removed from our crystallization construct (Fig. 1a). We found little contribution from these



**Figure 2 | Domain-domain contacts of HNF-4 $\alpha$ .** **a**, Circle indicates the convergence centre, where four domains (both LBDs, the upstream DBD and the downstream hinge) come together. **b**, Arg 91, located on the surface of the DBD, inserts deeply into a pocket at the base of the LBD-LBD surface. Numbers with prime refer to the upstream-positioned subunit. **c**, Ser 78, positioned on the back side of the DNA recognition helix, also enables the convergence of the DBD and the LBDs.

segments to the overall affinity of the complex for DNA, when examined individually or in combination (Fig. 1d). The proteolytically sensitive nature of these regions, even in the DNA-binding complex of HNF-4 $\alpha$ , also suggests that they are poorly ordered and not involved in DNA binding (Supplementary Fig. 1a-c). We also prepared the isolated AB and F domain fragments of HNF-4 $\alpha$  and tested their ability to bind to the rest of the homodimer-DNA complex. However, we detected no appreciable binding of these receptor portions with the rest of the complex (Supplementary Fig. 1d).

Both HNF-4 $\alpha$  LBDs have electron density profiles consistent with a trapped fatty acid, with the size indicating a myristic acid derived from *Escherichia coli*, in which the protein was expressed (Supplementary Fig. 2)<sup>15,16</sup>. The fatty acid is believed to lend structural integrity to the HNF-4 $\alpha/\gamma$  subfamily. Linoleic acid has been shown to be an exchangeable and potentially endogenous ligand of HNF-4 $\alpha$ , although this molecule does not confer transcriptional activity<sup>17</sup>. A stabilizing fatty acid, or a silent molecule that cannot switch receptor activity on and off, raises the question of how HNF-4 $\alpha$  activity is otherwise regulated.

The activities of NRs can be regulated by a variety of post-translational modifications (PTMs)<sup>18</sup>. In the case of HNF-4 $\alpha$ , two PTMs are well described for their ability to regulate receptor properties<sup>19,20</sup>. These modifications control the receptor's ability to bind DNA, and by extension its ability to regulate gene expression. We identified the quaternary sites of these PTMs within HNF-4 $\alpha$ . The first site, Arg 91, is a target of PRMT1, an enzyme that adds up to two methyl groups to the arginine side chain. Arg 91 methylation produces a marked enhancement in the DNA-binding activity of HNF-4 $\alpha$ <sup>19</sup>. The second site, Ser 78, is phosphorylated by protein kinase C (PKC), which disrupts the ability of HNF-4 $\alpha$  to bind DNA<sup>20</sup>. Therefore, taken together, these two PTMs act as on and off switches for regulating receptor activity.

Arg 91 methylation substantially enhances DNA affinity, but is not positioned to influence DNA binding directly from its location on the DBD farthest from the DNA. Figure 2b shows how its side chain deeply protrudes into the LBD-LBD cooperating surface that we described earlier as the receptor's multidomain convergence zone. There is a cavity directly above the side chain of Arg 91 to accommodate the two extra methyl groups, and the extension of the side chain through



methylation would more firmly ‘glue’ the DBD junctional interface with both LBDs. Therefore, this PTM acts to bias allosterically the receptor to bind DNA, by stabilizing the interdomain junctions associated with the final productive DNA complex.

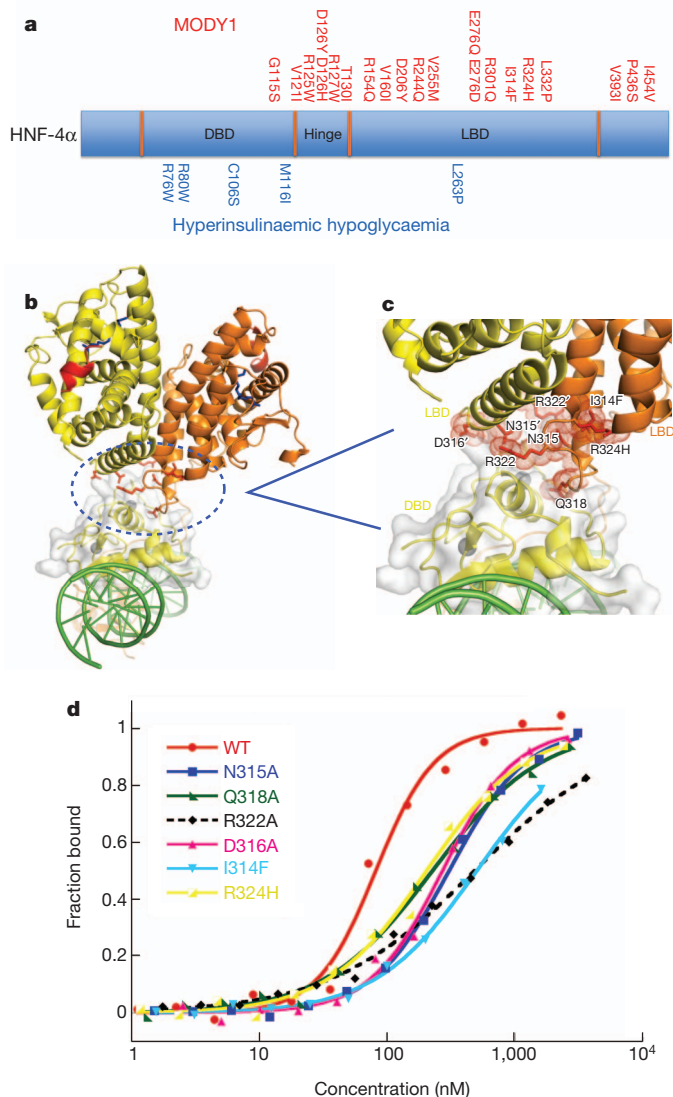
We next analysed the location of Ser 78, the site of PKC phosphorylation in a number of NRs<sup>20</sup>. Along with HNF-4 $\alpha$ , other NRs including FXR (also known as NR1H4), RAR- $\alpha$  (also known as NR1B1), VDR (also known as NR1I1), PPAR- $\alpha$  (also known as NR1C1), PXR (also known as NR1I2) and TR2 (also known as NR2C1) are similarly targeted by PKC, which in each case phosphorylates a similarly positioned serine on the DBD. Curiously, this serine always resides on the ‘wrong side’ of the DNA recognition helix, as is the case in HNF-4 $\alpha$ , where it seemingly cannot participate directly in DNA binding (Fig. 2c)<sup>20,21</sup>. Yet Ser 78 phosphorylation nevertheless weakens receptor–DNA binding substantially<sup>20</sup>. Our structure indicates that Ser 78 is positioned to engage the receptor’s interfacial connections so as to reduce DNA binding allosterically. Figure 2c suggests how an added phosphate group on this residue would create clash, both in

size and charge, with nearby Tyr 319, a residue that physically connects the receptor LBD with the DBD through Ser78. Phosphorylation would compromise the integrity of the quaternary fold needed for efficient DNA binding. Allosteric mechanisms of this type cannot be understood using the isolated crystal structures of DBDs or LBDs alone, as both Arg91 and Ser78 would appear to be too far from DNA-binding surface from that analysis. The current analysis of the domain organization, however, shows the unique positioning of these residues being consistent with their ability to affect the receptor's DNA-binding function allosterically.

We next asked whether some point mutations linked to MODY1 and hyperinsulinaemic hypoglycaemia are similarly positioned in sensitive interjunctional surfaces (Fig. 3a). For R76W and R80W mutations (associated with hyperinsulinaemic hypoglycaemia), there is a simple explanation for receptor dysfunction, as this pair of arginine residues directly contacts the AGGTCA half-sites (Supplementary Figs 9 and 10). V255M alters a residue that points into the LBD pocket, the only residue doing so among all the mutations associated with MODY1 and hyperinsulinaemic hypoglycaemia mutations. We found a number of mutations that lie at the sensitive domain–domain junctions of the complex. Sites such as R127W, D126Y, D126H and R125W locate to the downstream hinge region where they form domain–domain arrangements with the upstream DBD (Supplementary Figs 4 and 10). Mutational changes in this hinge site would misalign the interaction between domain–domain surfaces required to bridge the two DBDs into register with their successive AGGTCA half-sites. Indeed, we find that these mutant proteins substantially compromised DNA affinity (Supplementary Fig. 10). This loss of DNA binding also translates to a reduction in transcriptional activity<sup>22</sup>. We next examined MODY1 mutations I314F and R324H, and their adjacent residues (R322A, Q318A, D316A and N315A), which were found to be on the LBD and at the multidomain convergence centre of the complex (Fig. 3b–d). These mutations reduced the DNA affinity and transcriptional activity of the receptor (Fig. 3d and Supplementary Fig. 11).

Our investigation of PTMs and MODY1 mutations shows that changes introduced in the LBD, the hinge region, or in the DBD away from the DNA interface, still affect the DNA-binding properties of the receptor at a distance, by communicating through the interdomain junctions of the quaternary fold. It is interesting to note the subtlety of a single PTM or a single amino-acid mutational change, and the large distance with which these signals travel across the polypeptide to modulate DNA binding. Therefore, the domain convergence centre should be seen as both a sensitive centre for receiving signals, and an allosteric transmission system for propagating signals. At the same time, it is important to note that the two subunits of the homodimer are in altogether different environments owing to the asymmetric nature of the two subunits. PTM sites such as Ser 78 or Arg 91 will have a large influence on the complex only if they occur in the upstream DBD. In the same way, some MODY1 mutations would appear to be damaging if located in one, but not the other, subunit of the homodimer. Owing to the considerable numbers of genes being actively controlled by HNF-4 $\alpha$  in the liver and pancreas, the loss of even a fractional population of functional homodimers caused by heterozygous mutations can cause disease.

As both the HNF-4 $\alpha$  homodimer and PPAR- $\gamma$ -RXR- $\alpha$  complexes target DR1, we asked if their quaternary architectures were related. The common DR1 is expected to establish a similar DBD-DBD spacing in these complexes. Figure 4a, b shows the PPAR- $\gamma$ -RXR- $\alpha$  heterodimer and the HNF-4 $\alpha$  homodimer in an identical way, based on the layout of their common DR1 sequences. Figure 4c shows the superposition of these complexes when their DR1 sequences are aligned to match. Indeed, the DBDs occupy nearly identical positions in both DR1 complexes. Nevertheless, the higher-order quaternary arrangements are distinct for these two complexes (Fig. 4). In HNF-4 $\alpha$ , the LBDs are biased towards the upstream DBD, whereas in the PPAR- $\gamma$ -RXR- $\alpha$



**Figure 3 | Disease-linked mutations in HNF-4 $\alpha$ .** **a**, Summary of MODY1 and hyperinsulinaemic hypoglycaemia point mutations identified clinically in human populations. **b**, **c**, The MODY1 mutations, in many cases (residues in red) map to the convergence centre of the receptor domains (blue circle in **b**). **d**, DNA-affinity measurements of the wild-type (WT) and mutant receptors, as described in Fig. 1e. See also Supplementary Fig. 10 for studies of other disease-linked mutations.



complex the LBDs are biased towards the downstream RXR- $\alpha$  DBD. Moreover, the PPAR- $\gamma$ -RXR- $\alpha$  complex has its own type of domain convergence centre, which is not identical to that in the HNF-4 $\alpha$  complex.

The structural comparison indicates that the DNA response element type is not the only driver of quaternary structure in NRs. Receptor organization seems to be highly dependent on the constellation of non-conserved amino acids on these LBD surfaces, and on the length and sequence of the hinge segments, which are unique to NR members. We also note that DNA recognition is not identical in these two complexes. PPAR- $\gamma$  uses its hinge region to recognize an additional six base-pair segments located upstream to the DR1 core element, establishing the polarity of subunits in that heterodimer. HNF-4 $\alpha$  subunits do not use their hinge regions for DNA recognition, nor do they contact sequences outside the core DR1.

Our crystallographic findings with both NR complexes do not support the notion of a 'common architecture' for full-length NRs<sup>23</sup>. Our findings also dispel the view that NR polypeptides are arrays of 'domains on a string', each of which confers its own independent function without physical and functional integration. The repertoire of quaternary structures in the NR family is likely to be diverse, even though both the DBDs and LBDs are conserved. This expectation stems from the fact that neither hinge regions nor LBD surface residues are conserved in the NR family, yet these features are the key drivers of quaternary folding. The multiple response element configurations used in the NR family are another driver of quaternary organization.

Mounting evidence points to the importance of interdomain communication in the NR family. For oestrogen receptors, the activities of ligands are influenced by the response elements, and DNA can also influence coactivator binding<sup>24</sup>. In the glucocorticoid receptor, small conformational changes in the DBD propagate across the receptor to influence the LBD, and in the androgen receptor there is also evidence of DBD-to-LBD communication<sup>25,26</sup>. Our findings reveal that PTMs can modulate the interdomain connections in the quaternary fold. It has been reported that certain PPAR- $\gamma$  ligands can selectively block the phosphorylation of PPAR- $\gamma$ 2, indicating communications between the LBD pocket and the site of phosphorylation<sup>27,28</sup>. Ser 273 in PPAR- $\gamma$  is positioned within a domain-domain junction of the PPAR- $\gamma$ -RXR- $\alpha$  complex (Supplementary Fig. 12). From its position, the phosphorylation state of Ser 273 can communicate with the PPAR- $\gamma$  ligand-binding pocket and with the DNA-reading heads of the PPAR- $\gamma$ -RXR- $\alpha$  heterodimer.

For HNF-4 $\alpha$ , small molecules directed at the sensitive interjunctional sites may prove to be beneficial for treating MODY1 patients in which the DNA-binding properties have been mutationally compromised. To

find these molecules, high-throughput screening efforts must target the complete architecture of this receptor and not just the isolated LBD. We have identified two locations in the quaternary structure of the HNF-4 $\alpha$  complex that appear to be accessible for the binding of small-molecule allosteric modulators (Supplementary Fig. 13). An expanded understanding of the physical connectivity between LBDs, DBDs and other domains in the NR family should expand and better guide the discovery of receptor modulators with therapeutic value.

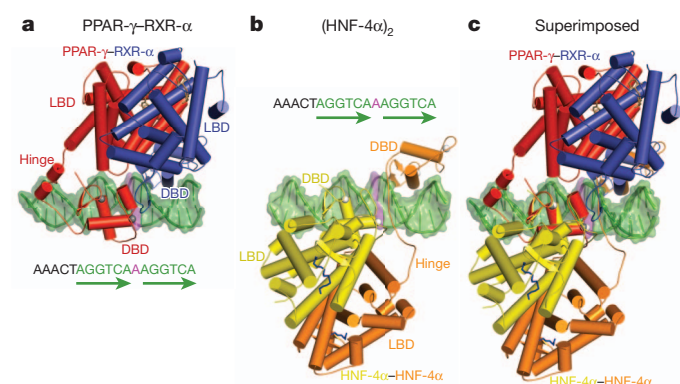
## METHODS SUMMARY

The HNF-4 $\alpha$  protein segment lacking the AB and F regions, corresponding to residues 46–368 (National Center for Biotechnology Information (NCBI) accession NM\_178849), was expressed in *E. coli*. The purified protein was then combined with synthetic duplex DNA (DR1) and coactivator peptide, further purified as a complex, and subjected to crystallization trials. X-ray diffraction data were collected at the Argonne National Laboratory SBC-CAT 19ID beamline and the structure was solved by molecular replacement, using the previously determined structures of the HNF-4 $\alpha$  LBD and DBD. Biochemical studies measuring DNA affinity were measured using wild-type and mutant proteins prepared using QuikChange Site-Directed Mutagenesis (Stratagene). The DNA was fluorescein labelled at the 5' end of the top strand, and fluorescence polarization anisotropy measurements were used to obtain  $K_d$  values.

**Full Methods** and any associated references are available in the online version of the paper.

Received 9 October 2012; accepted 28 January 2013.

Published online 13 March 2013.



**Figure 4 | Comparison the HNF-4 $\alpha$  homodimer and the PPAR- $\gamma$ -RXR- $\alpha$  heterodimer complexes on DR1 DNA. a**, The PPAR- $\gamma$ -RXR- $\alpha$  heterodimer on DR1. **b**, The HNF-4 $\alpha$  homodimer on DR1. **c**, Their overlap when the DR1 sequences are superimposed, showing the distinct domain-domain arrangements in these two complexes. The two complexes are shown in an identical fashion with respect to the DNA sequence facing the viewer.

- Sladek, F. M., Zhong, W. M., Lai, E. & Darnell, J. E. Jr. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev.* **4**, 2353–2365 (1990).
- Mangelsdorf, D. J. & Evans, R. M. The RXR heterodimers and orphan receptors. *Cell* **83**, 841–850 (1995).
- Bolotin, E. *et al.* Integrated approach for the identification of human hepatocyte nuclear factor 4 $\alpha$  target genes using protein binding microarrays. *Hepatology* **51**, 642–653 (2010).
- Wallerman, O. *et al.* Molecular interactions between HNF4a, FOXA2 and GABP identified at regulatory DNA elements through ChIP-sequencing. *Nucleic Acids Res.* **37**, 7498–7508 (2009).
- Yoon, J. C. *et al.* Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature* **413**, 131–138 (2001).
- Fang, B., Mane-Padros, D., Bolotin, E., Jiang, T. & Sladek, F. M. Identification of a binding motif specific to HNF4 by comparative analysis of multiple nuclear receptors. *Nucleic Acids Res.* **40**, 5343–5356 (2012).
- Odom, D. T. *et al.* Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**, 1378–1381 (2004).
- Ryffel, G. U. Mutations in the human genes encoding the transcription factors of the hepatocyte nuclear factor (HNF1) and HNF4 families: functional and pathological consequences. *J. Mol. Endocrinol.* **27**, 11–29 (2001).
- Ellard, S. & Colclough, K. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha (HNF1A) and 4 alpha (HNF4A) in maturity-onset diabetes of the young. *Hum. Mutat.* **27**, 854–869 (2006).
- Kapoor, R. R. *et al.* Hyperinsulinaemic hypoglycaemia. *Arch. Dis. Child.* **94**, 450–457 (2009).
- Flanagan, S. E. *et al.* Diazoxide-responsive hyperinsulinemic hypoglycemia caused by HNF4A gene mutations. *Eur. J. Endocrinol.* **162**, 987–992 (2010).
- Chandra, V. *et al.* Structure of the intact PPAR- $\gamma$ -RXR- $\alpha$  nuclear receptor complex on DNA. *Nature* **456**, 350–356 (2008).
- Nielsen, R. *et al.* Genome-wide profiling of PPAR $\gamma$ :RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes Dev.* **22**, 2953–2967 (2008).
- Jiang, G., Lee, U. & Sladek, F. M. Proposed mechanism for the stabilization of nuclear receptor DNA binding via protein dimerization. *Mol. Cell. Biol.* **17**, 6546–6554 (1997).
- Wisely, G. B. *et al.* Hepatocyte nuclear factor 4 is a transcription factor that constitutively binds fatty acids. *Structure* **10**, 1225–1234 (2002).
- Dhe-Paganon, S., Duda, K., Iwamoto, M., Chi, Y. I. & Shoelson, S. E. Crystal structure of the HNF4 $\alpha$  ligand binding domain in complex with endogenous fatty acid ligand. *J. Biol. Chem.* **277**, 37973–37976 (2002).
- Yuan, X. *et al.* Identification of an endogenous ligand bound to a native orphan nuclear receptor. *PLoS ONE* **4**, e5609 (2009).
- Weigel, N. L. & Moore, N. L. Steroid receptor phosphorylation: a key modulator of multiple receptor functions. *Mol. Endocrinol.* **21**, 2311–2319 (2007).
- Barrero, M. J. & Malik, S. Two functional modes of a nuclear receptor-recruited arginine methyltransferase in transcriptional activation. *Mol. Cell* **24**, 233–243 (2006).
- Sun, K. *et al.* Phosphorylation of a conserved serine in the deoxyribonucleic acid binding domain of nuclear receptors alters intracellular localization. *Mol. Endocrinol.* **21**, 1297–1311 (2007).

21. Gineste, R. *et al.* Phosphorylation of farnesoid X receptor by protein kinase C promotes its transcriptional activity. *Mol. Endocrinol.* **22**, 2433–2447 (2008).
22. Lu, P. *et al.* Structural basis of natural promoter recognition by a unique nuclear receptor, HNF4 $\alpha$ . *J. Biol. Chem.* **283**, 33685–33697 (2008).
23. Rochel, N. *et al.* Common architecture of nuclear receptor heterodimers on DNA direct repeat elements with different spacings. *Nature Struct. Mol. Biol.* **18**, 564–570 (2011).
24. Hall, J. M., McDonnell, D. P. & Korach, K. S. Allosteric regulation of estrogen receptor structure, function, and coactivator recruitment by different estrogen response elements. *Mol. Endocrinol.* **16**, 469–486 (2002).
25. Meijsing, S. H. *et al.* DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* **324**, 407–410 (2009).
26. Helsen, C. *et al.* Evidence for DNA-binding domain–ligand-binding domain communications in the androgen receptor. *Mol. Cell. Biol.* **32**, 3033–3043 (2012).
27. Choi, J. H. *et al.* Anti-diabetic drugs inhibit obesity-linked phosphorylation of PPAR $\gamma$  by Cdk5. *Nature* **466**, 451–456 (2010).
28. Choi, J. H. *et al.* Antidiabetic actions of a non-agonist PPAR $\gamma$  ligand blocking Cdk5-mediated phosphorylation. *Nature* **477**, 477–481 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by National Institutes of Health grants R01 DK094147 and R01 DK097475.

**Author Contributions** V.C. expressed, purified and crystallized the complex. P.H., along with V.C., solved and refined the structure and carried out the mutational binding studies. Y.K. collected, processed and reduced the X-ray diffraction data, and assisted with the molecular replacement search in structure determination. N.P., along with V.C., made the expression constructs for crystallization and for DNA-binding studies. D.W. carried out the transcription assays. F.R. supervised the work and wrote the manuscript.

**Author Information** Data have been deposited in the Protein Data Bank under accession 4IQR. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.R. ([frastinejad@sanfordburnham.org](mailto:frastinejad@sanfordburnham.org)).

## METHODS

**Expression, purification and crystallization.** The HNF-4 $\alpha$  proteins used in this study reference the NCBI sequence for HNF4, accession NM\_178849. All HNF-4 $\alpha$  (human) constructs, including FL (1–464),  $\Delta$ AB $\Delta$ F (46–368),  $\Delta$ F (1–368),  $\Delta$ AB (46–464) and DBD (46–126) and others were expressed from pET46 Ek/LIC vector in BL21(DE3) *E. coli* cells (Novagen). The HNF-4 $\alpha$   $\Delta$ AB $\Delta$ F (human) construct was used in crystallization experiments. Cells were induced with 0.5 mM IPTG at 17 °C for 16 h, and lysed in 20 mM Tris (pH 7.5), 500 mM NaCl, 20 mM imidazole and 10% glycerol. The purification used involved His-Bind resin (Novagen), SP-Sepharose column (GE Healthcare) and gel filtration (Superdex 200). Purified HNF-4 $\alpha$  was then combined with oligonucleotide (DR1) in a 1:1.5 molar ratio. The oligonucleotide strands 5'-GGAAGTAAAGGTCAG-3' and 5'-CCTGACCTTTGACCTAGTTC-3' were purified and annealed. Coactivator peptide (EKHKILHRLQDSY) was also added in a 3 $\times$  molar ratio. A final gel-filtration step was carried out to remove any excess DNA. Crystals were grown with 5–10% PEG 3350, 25 mM MgCl<sub>2</sub>, 25 mM NH<sub>4</sub>Cl, 10 mM dithiothreitol (DTT) and 0.1 M MES pH 6.5 at 4 °C.

**Data collection and structure determination.** Diffraction data was collected at the Argonne National Laboratory SBC-CAT 19ID beamline and the structure was solved by molecular replacement. X-ray data were collected at a wavelength of 0.9793 Å, at 100 K. The backbone dihedral (Ramachandran) angles for the amino acids in the final coordinates conform to preferred, allowed and disallowed statistics of 90.5%, 8.2% and 1.3%, respectively. The details of X-ray diffraction data collection, structure solution and refinement are in Supplementary Table 1.

**Fluorescence polarization assay.** All the mutants used in DNA-binding assays were prepared using QuikChange Site-Directed Mutagenesis (Stratagene). The DNA strands 5'-GGAAGTAAAGGTCAG-3' and 5'-CCTGACCTTTGACCTAGTTC-3' were annealed to make the DR1 for binding studies. For binding assays, 2 nM fluoresceinated DNA (5'-end conjugation on the top strand) was incubated with purified HNF-4 $\alpha$  protein for 2 h at room temperature (23 °C). Protein concentration was varied by serial dilution in binding buffer (20 mM

Tris-HCl, pH 7.8, 20 mM NaCl, 8% glycerol, 10 mM DTT). The fluorescence polarization signals were recorded using 96-well black polystyrene plates on FlexStation 3. The data were later converted to fluorescence anisotropy values and normalized. The  $K_d$  of each construct and mutants were calculated by fitting the curve in KaleidaGraph 4.1.

**Interactions of the A/B and F domains within the HNF-4 $\alpha$  complex.** For the study shown in Supplementary Fig. 1D, the AB domain (residues 1–45) and F domain (369–465) of human HNF-4 $\alpha$  were expressed from pET46 Ek/LIC vector in BL21(DE3) *E. coli* cells (Novagen). Each protein was purified using a HisTrap column, followed by size-exclusion chromatography on a HiLoad 16/60 Superdex 200 column. The final samples were prepared in 10 mM phosphate (pH 7.0) and 100 mM NaCl buffer for fluorescein labelling. Ten micrograms of each peptide, HNF-4 $\alpha$ -AB, HNF-4 $\alpha$ -F and PGC-1 $\alpha$  LXXLL-motif peptide (AEEPSLLKKLLAY, synthetically made and purchased from AnaSpec) were incubated with fluorescein-5-EX succinimidyl ester at a molar ratio of 1:2 in 100 mM potassium phosphate (pH 7.0) coupling buffer at 37 °C for 60 min, quenched by adding 100 mM Tris (pH 8.0) and incubated for 30 min at room temperature. The labelled peptides were purified using Sephadex G-15 column for fluorescence polarization measurements. The HNF-4 $\alpha$  $\Delta$ AB $\Delta$ F–DNA complex was prepared for the peptide-binding assays as described earlier. Five nanomolar of each fluorescein-labelled peptide was incubated with HNF-4 $\alpha$ –DNA complex of serially diluted concentrations for 2 h at room temperature. Recording of fluorescence polarization signals and data processing were conducted as described earlier.

**Transcription reporter assays.** For the transcriptional reporter studies shown in Supplementary Fig. 11, we used both HEK293T and COS-7 cells that were seeded in 24-well plates and 1 day later transfected with 400 ng of the pCMV-Tag1-HNF-4 $\alpha$  (46–368) wild-type or mutant plasmid, 100 ng of apoCIII-pTKLuc reporter (a gift from D. Kelly) and 10 ng of pRL (control *Renilla* luciferase) using jetPEI reagent (Polyplus) according to the manufacturer's protocol. Luciferase activity was measured 48 h after transfection using the Dual-Glo Luciferase Assay System (Promega) and data were normalized by the relative ratio of firefly and *Renilla* luciferase activity.

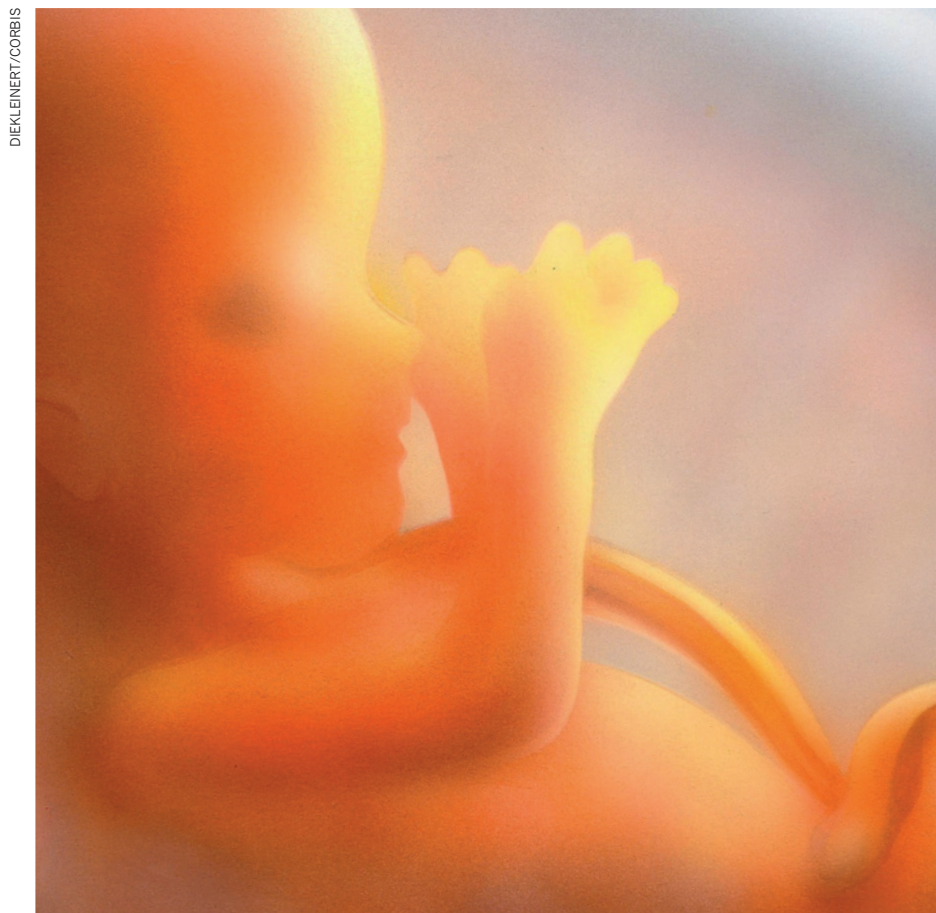


# CAREERS

**TURNING POINT** Work on genetics of twinning leads to award and post **p.401**

**CAREERS TOOLKIT** Career advice and support for researchers [go.nature.com/mfgmok](http://go.nature.com/mfgmok)

**NATUREJOBS** For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)



DIETKLEINERT/CORBIS

## REPRODUCTIVE BIOLOGY

# Breeding opportunities

*The reproductive sciences offer an unusual variety of career options — but some are more fertile than others.*

BY CHARLOTTE SCHUBERT

**H**elp to implant a human embryo; watch mouse eggs divide after fertilization; interview women about their experiences with emergency contraception; collect sperm from finches in the Galapagos Islands off Ecuador: specialists in

reproductive biology undertake these duties and many more. The field provides a surfeit of career trajectories and research questions. “We cover everything from fertilization to death,” says Dolores Lamb, director of the Center for Reproductive Medicine at Baylor College of Medicine in Houston, Texas.

The reproductive sciences also touch on

human life and biology at almost every level, from molecular and cellular events such as the recognition of the egg by sperm, to whole-body processes including hormonal regulation of puberty and population-level questions such as what factors affect teenage-pregnancy rates.

“It’s amazing, the diverse backgrounds that all somehow feed into reproductive medicine,” says Lamb. Many researchers study reproduction as part of a doctorate in reproductive or developmental biology. Others might find their way to the field through cell biology (focusing on sperm stem cells, for example) or animal sciences (perhaps studying cattle hormones). Attendees at the annual meeting of the Society for the Study of Reproduction, which is based in Madison, Wisconsin, hail from a wide range of disciplines.

Diversity does not guarantee jobs, however, and positions are scarce in many areas of industry and academia. Richard Sharpe, who heads the graduate programme at the UK Medical Research Council’s Centre for Reproductive Health at the University of Edinburgh, says that three years ago he assured students that if they excelled, they could find work. “We can no longer say that,” he says. Academic opportunities have shrunk as the recession has taken a bite out of budgets and funding priorities have shifted to areas such as chronic disease, says Sharpe. (Although location does matter; see ‘Renminbi for reproduction’.) In the reproductive sciences, as in most life sciences, the job market has constricted, agrees Michael Skinner, founder of the Center for Reproductive Biology at Washington State University (WSU) in Pullman.

Even so, he says, studying reproduction will give doctoral students an edge with employers from fertility clinics to animal-agriculture companies. And areas such as global health continue to grow.

## LOOKING TO ACADEMIA

WSU graduates have landed academic jobs in everything from toxicology to oncology, says Skinner, who notes that much cancer research focuses on reproduction-related cancers such as breast and prostate. These days, “you have to market yourself broadly”, says Tracy Clement, who earned her PhD with Skinner and is looking for an academic post. “It does not feel like a good time to be on the job market.”

If she does get a university job, Clement will face funding difficulties. The major source of US research money in the field is the National Institute of Child Health and Human Health. ►

► Development in Bethesda, Maryland, but last year only 12.5% of grant applications for major research projects to the institute were successful — less than at many other National Institutes of Health (NIH) institutions.

In the future, says Skinner, academics will require support from multiple sources — something for which the broad field is particularly suited. Skinner, for instance, has received funding not only from the NIH, but also from the Bill & Melinda Gates Foundation in Seattle, Washington, to develop a male contraceptive and from the US Department of Defense to study how exposure to environmental toxicants affects subsequent generations. He has also applied to the John Templeton Foundation, based in West Conshohocken, Pennsylvania, to study the role of epigenetics in finch evolution in the Galapagos.

### MIXED OUTLOOK IN INDUSTRY

The pharmaceutical industry has shed jobs in contraception over the past 10 years, in part because of nervousness about the side effects of reproduction-related drugs, which have drawn numerous lawsuits. Companies such as Wyeth — acquired in 2009 by Pfizer in New York — and Bayer, based in Leverkusen, Germany, have cut back or dropped entire research programmes.

Daniel Johnston, a former principal research scientist at Pfizer who was laid off in 2010, could not find a job directly related to human reproduction at a pharmaceutical or biotechnology company despite more than 5 years of pharmaceutical experience in contraception and women's health. But while at Pfizer, Johnston had begun to work in oncology — and that experience helped him to find work at a company focused on liver cancer.

Johnston's lateral move is not unusual, given that reproduction overlaps with many areas (Johnston compares a testis to a tumour; both have a low-oxygen core of rapidly dividing stem cells — but one produces sperm instead of cancer cells). Susan Fisher, director of translational research in perinatal biology and medicine at the Center for Reproductive Sciences of the University of California, San Francisco, says that adaptability gives her graduates a leg up in industry. They have found jobs at biotechnology companies focused on oncology, stem cells and the rapidly expanding area of prenatal genetic testing, a market that could soon be worth more than US\$1 billion yearly (see *Nature* **486**, 454; 2012). Companies competing in this area include Ariosa Diagnostics in San Jose and Natera in San Carlos, both in California. Skinner advises students and postdocs interested in an industry job to get training in the broadest possible range of lab technologies, including genomics.

Reproductive biologists are also finding work in animal agriculture. "I get calls maybe two or three times a year from companies looking for a master's- or PhD-level scientist

to run a lab," says Derek McLean, a biologist who studies cattle, pig and mouse reproduction at WSU.

Scientists who can store and manage animal semen and test fertility products for female livestock are in demand at companies and organizations such as Select Sires, a federation of farmer cooperatives based in Plain City, Ohio, that provides livestock-breeding services.



*"You have to prepare for option one — but have option two and three in the wings."*

Tracy Clement

for such posts have generally trained in a reproductive-biology laboratory and have a degree in animal science. The job market is steady in the United States, and is expanding in Brazil and other emerging countries that are moving towards a more industrialized animal-agriculture system, says McLean, who is leaving WSU this autumn for Phibro Animal Health in Teaneck, New Jersey.

Genomics opportunities are also emerging at companies that are developing tests for genetic selection of farm animals; one such employer is Illumina in San Diego, California, which markets a high-density Bovine BeadChip, a genetic array that detects traits in cattle.

Fertility clinics are an option for those looking for something more applied — and perhaps more likely to have an immediate impact on

people's lives. There are about 400 clinics in the United States, and they often hire PhD-level scientists to manage a staff of bachelor's- and master's-level technicians involved in tasks such as culturing and freezing human embryos and performing hormone assays. Such scientists often interact with patients — a lead embryologist, for instance, might contact patients with test results, and assist during implantation of the embryo. "It brings a lot of joy in our profession to help people to have a family," says Pierre Miron, who heads a fertility clinic near Montreal in Canada. "The interaction with patients is a great part of the job." Jobs at clinics affiliated with universities often provide more opportunities for fertility research than private clinics, says Miron.

The government of Quebec province began paying for *in vitro* fertilization in 2010 as part of the national health-care system, leading to rising demand for services. Many clinics are filling job slots with trainees from Réseau Québécois en reproduction (the Quebec Reproduction Network), a consortium of more than 80 researchers at institutions such as McGill University in Montreal. But in most of the world, jobs in fertility clinics are competitive. Strong candidates not only have the right human touch, says Miron, but also have training in a range of techniques, including sperm and egg manipulation.

### GLOBAL IMPACT

Global health may be a more fruitful area for reproductive-biology graduates. "My personal feeling is that there is now a stronger market for people working at the population level than at the molecular level in the reproductive sciences," says Ward Cates, president emeritus of FHI 360, a global-health organization based

## ASIAN OUTLOOK

### Renminbi for reproduction

Thirteen years after earning his PhD, and following two stints as a postdoc, Minghan Tong, a reproductive biologist at Washington State University in Pullman, has finally landed a tenure-track research position. Unable to find such a job in the United States, Tong is heading to greener pastures in his native China. In September, he will begin work at the Shanghai Institutes for Biological Sciences. Tong and others are applying for a 39-million-renminbi (US\$6.3-million) grant to study the epigenetic regulation of sperm production.

China is a bright spot in the tight market for academic jobs in reproductive sciences. The sperm-production grant is part of a major funding initiative in the field by the Chinese government, which is interested in new approaches to contraception and in

the reproductive impacts of environmental problems such as toxicants that may damage sperm and eggs by tweaking the epigenome to affect multiple generations.

In 2007, the Chinese Ministry of Science and Technology designated development and reproductive sciences as one of four core research areas, ramping up grant funding. This year, the ministry is expected to fund four 5-year projects in the field, each worth more than 24 million renminbi.

Other national and local agencies have also increased funding, spurring the establishment of dozens of research centres focused on reproductive sciences and creating hundreds of jobs, says Qinghua Shi, a professor of life sciences at the University of Science and Technology of China in Hefei. **C.S.**



in Durham, North Carolina. He points to Family Planning 2020, an ambitious initiative to roll out contraceptive services to 120 million girls and women in developing countries by 2020. Donors including the Bill & Melinda Gates Foundation and governments of both developed and developing nations have pledged \$2.6 billion to the programme, which was launched last July at a meeting in London spearheaded by Melinda Gates. “This will define the future for public-health jobs in the reproductive sciences,” says Cates.

He adds that the initiative will create jobs, mostly in the developing world, for researchers who know how to cost-effectively implement such services and for scientists who can evaluate their impact — by, for instance, assessing the uptake of contraception and its effects on population growth and women’s and children’s health. The effort will require researchers with backgrounds in areas such as demography, sociology, economics and public health.

Cates says that researchers with a basic-science background in reproductive sciences and extra training in fields such as epidemiology often have a leg up when competing for jobs in areas including clinical-trial design, because of their understanding of biology.

Patricia Sadate-Ngatchou earned a PhD studying sperm development at WSU. But a visit home to Cameroon during a major cholera outbreak in 2010 changed the course of her career. “How do you help people on the ground?” she asked herself.

Sadate-Ngatchou is now studying for a master’s degree in epidemiology at the University of Washington in Seattle. Her ultimate goal is to move into a decision-making position in government or a foundation involved in reproductive health; a suitable post might be as a programme officer overseeing grants. However, Sadate-Ngatchou thinks that she may first have to do entry-level work as an epidemiologist, for instance in disease surveillance.

The variety of questions and opportunities in reproductive biology keeps some researchers hooked on the field, despite the tough market. Some end up in niches they never expected, such as facilitating panda or reptile reproduction in zoos, or assessing toxicants for their effects on embryonic and pubertal development at government institutions such as the US Environmental Protection Agency. Clement is open to a variety of possibilities. “If you are a reproductive biologist,” she says, “you have to prepare for option one — but have option two and three in the wings.” ■

**Charlotte Schubert** is a freelance writer based in Seattle, Washington.

## TURNING POINT

# Bruno Reversade

*A molecular biologist at the Agency for Science, Technology and Research (A\*STAR) Institute of Medical Biology in Singapore, Bruno Reversade is the first scientist based outside Europe to win the European Molecular Biology Organization Young Investigatorship Award, which he collected last November for his work on genetics and twinning.*

### Do you thrive in a competitive environment?

Yes. I realized that university would be competitive when one of my first professors said that 60% of the class would not make it to the second year. After that, I sat in the front and worked hard.

### What led to your fascination with embryonic development?

I went to the University of Western Ontario in London, Canada, in my fourth year as an undergraduate and worked on early zebrafish development. The developing embryo was so beautiful and fundamental to life that I realized it was a special area. At the time, medicine and biology were all about identifying and treating disease, but I found a resonance with birth and development. I went to the Pasteur Institute in Paris for a year to work on early head development after seeing a knockout mouse with no head on the cover of *Nature* (W. Shawlot and R. R. Behringer *Nature* 374, 425–430; 1994).

### What was the biggest challenge of your PhD?

I did most of my research at the University of California, Los Angeles, where I spent the first half of my programme chasing chordin, a protein that my adviser, Eddy De Robertis, and I thought was circulating in the blood. After three years we found we were mistaken. I persevered, however, and we detailed how multiple proteins help embryos that are cut in two to self-regulate consistently. Eddy and I published the work in *Cell* (B. Reversade and E. M. De Robertis *Cell* 123, 1147–1160; 2005).

### What led you to focus on genetics in twinning?

Sitting in the lab cutting frog embryos day after day led me to a defining realization. Identical twins occur once in every 300 births, more frequently than most genetic diseases. The dogma at the time was that twinning just happens, but I started to look for evidence of a genetic trigger. Then Hanan Hamamy, a genetic clinician who at the time was at the Jordanian National Center for Diabetes, Endocrinology and Genetics in Amman, identified 13 pairs of identical twins across multiple generations of a single family — hinting at a genetic link.



She kindly invited me to work with her.

### How were you able to rush out to Jordan?

I wanted to test my ideas as quickly as possible, and I didn’t want to do a postdoc. Perhaps that was arrogant or unrealistic, but I wanted to be independent. Fortunately, Swiss philanthropist Branco Weiss was seeking young scientists who were pursuing a biological problem with societal impact for the Society in Science fellowship. I met Branco and explained that identical twins can develop through several mechanisms, including embryonic bisection and possibly genetics, which for me calls into question the moral uproar over cloning. I convinced him that the idea was worth pursuing and he gave me the money.

### Have you published this work?

Not yet. I’m now working with samples from other families with multiple sets of twins. We found a gene that is overexpressed in identical twins and encodes a protein. We are making sure it is well protected by patents.

### What is your most important career move so far?

In 2008, I was the first A\*STAR investigator recruited as an assistant professor at the Institute of Medical Biology. They offered me carte blanche: I have no teaching or grant-writing responsibilities. Everything was new and the country was investing so much in science. I have blossomed here because I got that freedom just as the revolution in human genetics began.

### What do you plan to do next?

I want to work on rare diseases ranging from developmental anomalies to inherited cancers. If you want to understand a trait in the general population, you need to look at the outliers. ■

INTERVIEW BY VIRGINIA GEWIN



# THE SEPARATISTS

*A radical element.*

BY KJ KABZA

The last thing I remembered was the Hot chair and the grey CAT-15 cable jamming into the jack socket in my skull. I knew what that meant.

I rolled over in bed. My new, standardized body smelled of the stasis tank and moved with clumsy naivety.

"Car bomb," said Sara.

I looked up. She stood at my bedside like a coroner, gazing down impassively at another piece of meat. Her tone was flat and her face indifferent stone. "Outside the Smithsonian. The Separatists again."

I turned away from her dead eyes. Even after all this, her loathing is what kills me. I know my continuity is a myth, and I'm nothing but an echo. But her eyes never say, *At least I've got the echo.* They say, *You are a monster — a pre-made flesh lattice, stamped with my dead husband's face.*

"Oh," I said.

We waited in silence for the doctor.

In 15 days I was back above ground, dressed in a dark grey suit and a red power tie. Smiling and confident. Tone warm yet brisk. Sara at my side, one hand on my arm, smiling, always smiling. How long could we go on like this? *Mr President, how are you feeling? Mr President, are you going to reopen talks with Alabama? Mr President, how do you respond to allegations of terrorism against the Confederacy of Texahoma?*

While I was out, I missed the secession of South Missouri. The Secretary of Homeland Integrity said they'd left nine days ago, issuing a statement with the by-now familiar elements: time to leave a once-great nation, fallen into extreme financial and social decay, too divided to fight for what we needed. Time for something radical. "And by the way," said the secretary, "South Missouri's so-called Prime Minister is agitating for war."

*In times like these,* Dr Ramaran had said, after dubbing my makeshift immortality 'Operation Phoenix', *the President has to seem untouchable.*

Sara came to bed one night while I was still awake. "Nick."

"Yes?"

Around us, the darkness rustled as she burrowed under the blankets. Her voice was distant, spoken to the wall at my back instead of me. "Before you died last time."



"Don't."

She paused. I knew I'd said the wrong thing, but how could I ever say anything else?

Sara went ahead. "My mother called to talk about Lacey. And... you and I sat in the garden, for a while."

She expected me to nod. So I did.

"And you complimented the President of Transnistria on her perfume."

I bit my tongue. She kept talking, placing the unimportant moments in a careful row, as if the order of their occurrence could give them any meaning. Like I always did, I pretended she was debriefing me on the loose ends accumulated since the last back-up, and not reminding me, again, in every possible way, of what she'd lost.

Her litany trickled into nothing. Empty minutes passed. We listened to each other's breaths, the only proof of ourselves in the darkness.

I said, "I'm sorry."

Colourlessly: "I know."

"If it were up to me —"

"It isn't."

"If it were up to me," I repeated, "I would've run away with you after the first time. Or maybe I just should've done it anyway, consequences be damned."

The blankets pulled as she sat up. "You mean the first one of you should've."

"Sara, what can I do?"

"You can bring my

husband back."

"Goddammit —"

"You can die," she said, "but a copy of a sixth-generation copy would just take your place. You could divorce me, but the Separatists would be all over that. You could even blow up that goddam laboratory. They still haven't admitted to mind transfers in humans with anyone other than you. How long are you planning on going on like this?"

"Or you know what? You could murder me. You might as well, for what I feel every time you —" Her voice broke. "— every time another one of you comes into my house."

"Sara —"

"I could've lived with it after the first time, Nick. When your first copy was as scared as I was, and neither one of us knew what it all meant, and he thought he wasn't the same person either. But I don't even know what you are, now."

"I know I'm not the same person. But —"

"You're not even close, Nick." Only then did she start to cry. "You're not close enough at all."

I got out of bed and left the room, the agents outside pretending that they didn't hear her sobs.

I went into the lab that night, alone. I'd say I didn't know what I was thinking, but that's a lie. I was thinking of Sara.

The stasis tanks awaited me, where hairless, humanoid lattices lay in indefinite repose. One was already prepped and imprinted with my face, and I set that tank to drain.

I laid the body on a bed, inserted the CAT-15 cable into the socket at the base of the skull, and accessed the Hot computer, where every back-up of my consciousness resides. I found the oldest conglomerate that was flagged as 'utilized' — the only one Sara had said she could live with — and prepared the download.

While I waited, I wrote out a note for him on a piece of scrap paper:

"Run away with Sara. It's time for something radical." ■

KJ Kabza's work has appeared in F&SF, Daily Science Fiction, Beneath Ceaseless Skies, AE: The Canadian Science Fiction Review and others. Visit him at [www.kjkabza.com](http://www.kjkabza.com) or follow him on Twitter @KJKabza.

JACEY

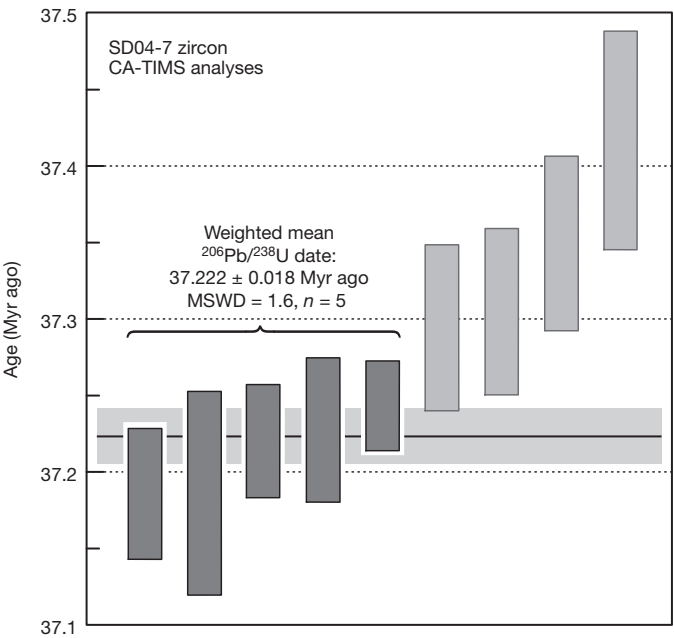
A Late Eocene date for Late Triassic bird tracks

ARISING FROM R. N. Melchor, S. De Valais & J. F. Genise *Nature* **417**, 936–938 (2002)

Bird-like tracks from northwest Argentina have been reported as being of Late Triassic age<sup>1</sup>. They were attributed to an unknown group of theropods showing some avian characters. However, we believe that these tracks are of Late Eocene age on the basis of a new weighted mean <sup>206</sup>Pb/<sup>238</sup>U date (isotope dilution–thermal ionization mass spectrometry method) on zircons from a tuff bed in the sedimentary succession containing the fossil tracks. In consequence, the mentioned tracks are assigned to birds and its occurrence matches the known fossil record of Aves.

The redbed sequence of the former Santo Domingo Formation yielded several-hundred bird-like footprints, which were assigned to *Gruipeda dominguensis* (the most common ichnotaxon), cf. *Alaripeda* isp., and another taxonomically indeterminate bird-like footprint<sup>2</sup>. The age of the stratigraphic unit was considered to be Late Triassic on the basis of known fossil wood remains and geochronological information from basalt lava flows thought to be interbedded in this unit<sup>1,3,4</sup>. Further geological studies revealed that the Santo Domingo Formation contains several thrust sheets of different ages, and that the trace-fossil-bearing horizons belong instead to the recently proposed Laguna Brava Formation, in a thrust sheet separate from the one that contains the dated basalt and fossil wood remains<sup>5</sup>. A 12-cm-thick crystal-rich ash-fall tuff within the thrust sheet with the bird-like footprints was sampled for this study for U–Pb zircon geochronology (supported by US National Science Foundation grant EAR 0931839 and ANPCyT PICT 13286 from Argentina). This tuff lies 38 m below the first layer with definite *G. dominguensis* and 124 m below the main horizon with hundreds of *G. dominguensis*<sup>6</sup>. There is no stratigraphic discontinuity between the tuff bed and the footprint-bearing levels. Zircon grains were separated from the tuff using conventional methods and were dated using high-precision chemical abrasion–isotope dilution–thermal ionization mass spectrometry (CA-ID-TIMS)<sup>7,8</sup>. The tuff yielded abundant clear, long bipyramidal (150 μm and 250 μm) and sharply faceted zircons. The five youngest analyses from a total of nine form a coherent cluster with a weighted-mean <sup>206</sup>Pb/<sup>238</sup>U date of 37.222 ± 0.018/0.024/0.047 million years (Myr) ago (internal uncertainties/with tracer calibration uncertainties/with decay constant uncertainties; mean square of weighted deviation = 1.6) (Fig. 1 and Table 1).

Therefore, we suggest that the maximum age for the bird tracks of the Laguna Brava Formation of northwest Argentina (including *G. dominguensis*) is Late Eocene (Bartonian/Priabonian<sup>9</sup>), and matches the known fossil record of birds. Recent studies<sup>10</sup> documenting possible



**Figure 1 | Date distribution plot for analysed zircons from tuff sample SD04-7 from the former Santo Domingo Formation (now considered as part of the Laguna Brava Formation).** The y axis shows measured <sup>206</sup>Pb/<sup>238</sup>U age (see Table 1 for complete analytical data). Data acquisition, reduction, error propagation and plotting were done using EARTHTIME supported statistical methods<sup>12</sup> and the U-Pb\_Redux software package<sup>13</sup>. Bar heights are 2σ analytical uncertainties; dark/light grey boxes signify analyses included/excluded from age calculation, respectively. The shaded horizontal band and its width represent the calculated weighted mean age and its 2σ analytical error. MSWD, mean square of weighted deviation; n, number of analyses included in the calculated date.

**Table 1 | U–Pb data for analysed zircon from tuff sample of the Laguna Brava Formation**

Sample fractions†	Composition				<sup>208</sup> PbII/ <sup>206</sup> Pb	Ratios						Age (Myr ago)				Corr. coef.
	Pb <sub>c</sub> ‡ (pg)	Pb*‡/ Pb <sub>c</sub>	Th/ U	<sup>206</sup> Pb§/ <sup>204</sup> Pb		<sup>206</sup> Pb¶/ <sup>238</sup> U	Error (2σ%)	<sup>207</sup> Pb¶/ <sup>235</sup> U	Error (2σ%)	<sup>207</sup> Pb¶/ <sup>206</sup> Pb	Error (2σ%)	<sup>206</sup> Pb/ <sup>238</sup> U	Error (2σ)	<sup>207</sup> Pb/ <sup>235</sup> U	<sup>207</sup> Pb/ <sup>206</sup> Pb	
SD04-7																
z2	0.8	7.8	0.90	441.2	0.287	0.005821	(0.19)	0.03811	(2.31)	0.04750	(2.27)	37.417	0.071	37.98	74	0.23
z3	1.4	21.1	0.74	1,211.7	0.238	0.005804	(0.15)	0.03765	(0.83)	0.04707	(0.81)	37.305	0.054	37.53	52	0.26
z4	0.7	21.8	0.65	1,281.5	0.207	0.005785	(0.12)	0.03728	(0.90)	0.04676	(0.87)	<b>37.186</b>	<b>0.043</b>	37.16	36	0.32
z5	0.5	48.2	0.67	2,792.1	0.214	0.005791	(0.10)	0.03753	(0.39)	0.04702	(0.36)	<b>37.221</b>	<b>0.037</b>	37.41	49.4	0.33
z7	0.5	22.1	0.62	1,305.7	0.198	0.005802	(0.15)	0.03752	(0.82)	0.04692	(0.80)	37.295	0.054	37.40	44	0.27
z8	1.3	11.2	0.59	673.2	0.190	0.005792	(0.13)	0.03747	(1.52)	0.04694	(1.49)	<b>37.228</b>	<b>0.047</b>	37.35	45	0.27
z9	0.7	26.5	0.62	1,561.8	0.199	0.005811	(0.15)	0.03737	(0.72)	0.04666	(0.69)	37.350	0.057	37.25	31	0.32
z10	1.6	13.7	0.73	791.2	0.234	0.005785	(0.18)	0.03727	(1.83)	0.04674	(1.73)	<b>37.187</b>	<b>0.066</b>	37.15	35	0.60
z11	0.5	46.6	0.59	2,756.2	0.189	0.005794	(0.08)	0.03728	(0.44)	0.04669	(0.41)	<b>37.244</b>	<b>0.029</b>	37.17	32.4	0.43

Corr. coef., correlation coefficient. Pb\* is radiogenic Pb concentration. Age calculations are based on the decay constants of ref. 11.  
†All analyses are single zircon grains and pre-treated by the thermal annealing and acid leaching (CA-TIMS) technique. Data used in age calculations are in bold.  
‡Pb<sub>c</sub> is total common Pb in analysis.  
§Measured ratio corrected for spike and fractionation only.  
||Radiogenic Pb ratio.  
¶Corrected for fractionation, spike, blank and initial Th/U disequilibrium in magma. All common Pb is assumed to be blank. Total procedural blank was less than 0.1 pg for U. Blank isotopic composition: <sup>206</sup>Pb/<sup>204</sup>Pb = 18.42 ± 0.35; <sup>207</sup>Pb/<sup>204</sup>Pb = 15.36 ± 0.23; <sup>208</sup>Pb/<sup>204</sup>Pb = 37.46 ± 0.74.

landing trace fossils and probing marks closely associated with *G. domingensis* strongly suggested a younger age for the host sedimentary succession. Re-definition of stratigraphic units in this region and further supporting geological and palaeomagnetic evidence have been published elsewhere<sup>5</sup>.

**Ricardo N. Melchor<sup>1</sup>, Robert Buchwaldt<sup>2</sup> & Samuel Bowring<sup>2</sup>**

<sup>1</sup>INCITAP (CONICET-UNLPam), Av. Uruguay 151, 6300 Santa Rosa, La Pampa, Argentina.

email: rmelchor@exactas.unlpam.edu.ar

<sup>2</sup>Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139-4307, USA.

Received 26 November 2012; accepted 9 January 2013.

1. Melchor, R. N., De Valais, S. & Genise, J. F. Bird-like fossil footprints from the Late Triassic. *Nature* **417**, 936–938 (2002).
2. De Valais, S. & Melchor, R. N. Ichnotaxonomy of bird-like footprints: An example from the Late Triassic-Early Jurassic of northwest Argentina. *J. Vertebr. Paleontol.* **28**, 145–159 (2008).
3. Caminos, R., Zamuner, A. B., Limarino, O. C. & Fauqué, L. El Triásico superior fosilífero en la Precordillera riojana. *Rev. Asoc. Geol. Argent.* **50**, 262–265 (1995).
4. Coughlin, T. J. *Linked Orogen-Oblique Fault Zones in the Central Argentine Andes: The Basis of a New Model for Andean Orogenesis and Metallogenesis* (Univ. Queensland, 2001).
5. Vizán, H. *et al.* Geological setting and paleomagnetism of the Eocene red beds of Laguna Brava Formation (Quebrada Santo Domingo, northwestern Argentina). *Tectonophysics* **583**, 105–123 (2013).
6. Melchor, R. N., Bedatou, E., de Valais, S. & Genise, J. F. Lithofacies distribution of invertebrate and vertebrate trace-fossil assemblages in an Early Mesozoic ephemeral fluvio-lacustrine system from Argentina: Implications for the *Scoyenia* ichnofacies. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **239**, 253–285 (2006).
7. Mattinson, J. M. Zircon U–Pb chemical abrasion (“CA-TIMS”) method: Combined annealing and multi-step partial dissolution analysis for improved precision and accuracy of zircon ages. *Chem. Geol.* **220**, 47–66 (2005).
8. Rioux, M. *et al.* Protracted timescales of lower crustal growth at the fast-spreading East Pacific Rise. *Nature Geosci.* **5**, 275–278 (2012).
9. Gradstein, F. M., Ogg, J. G., Smith, A. G., Bleeker, W. & Lourens, L. J. A new Geologic Time Scale, with special reference to Precambrian and Neogene. *Episodes* **27**, 83–100 (2004).
10. Genise, J. F. *et al.* Application of neoichnological studies to behavioural and taphonomic interpretation of fossil bird-like tracks from lacustrine settings: The Late Triassic-Early Jurassic? Santo Domingo Formation, Argentina. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **272**, 143–161 (2009).
11. Jaffey, A. H., Flynn, K. F., Glendenin, L. E., Bentley, W. C. & Essling, A. M. Precision measurement of half-lives and specific activities of <sup>235</sup>U and <sup>238</sup>U. *Phys. Rev. C* **4**, 1889–1906 (1971).
12. McLean, N. M., Bowring, J. F. & Bowring, S. A. An algorithm for U–Pb isotope dilution data reduction and uncertainty propagation. *Geochem. Geophys. Geosyst.* **12**, Q0AA18 (2011).
13. Bowring, J. F., McLean, N. M. & Bowring, S. A. Engineering cyber infrastructure for U–Pb geochronology: Tripoli and U–Pb\_Redux. *Geochem. Geophys. Geosyst.* **12**, Q0AA19 (2011).

**Author Contributions** R.N.M. conducted project planning, geological constraints and tuff sampling. R.B. and S.B. carried out laboratory work, data analysis and presentation.

**Competing Financial Interests** Declared none.

doi:10.1038/nature11931

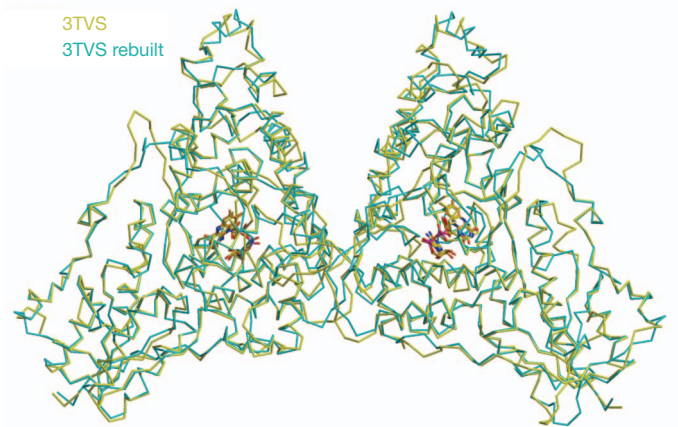


Updated structure of *Drosophila* cryptochrome

ARISING FROM B. D. Zoltowski *et al.* *Nature* **480**, 396–399 (2011)

Recently, we determined the X-ray crystal structure of full-length cryptochrome from *Drosophila*<sup>1</sup>. Here we report an improved model of the *Drosophila* cryptochrome (dCRY) structure that corrects errors in the original coordinates (Protein Data Bank (PDB) accession 3TVS). Further refinement of the structure, with automated rebuilding algorithms in Phenix<sup>2</sup> followed by manual building, indicated that a model of dCRY could be produced with excellent refinement statistics without taking into account the non-merohedral twinning originally reported (Table 1).

The rebuilt structure has a root mean squared deviation (r.m.s.d.) on Cα positions of 2.4 Å compared to the deposited coordinates, with most differences found in the conformation of surface loops (Fig. 1). However, the new analysis also indicates that the sequence register of the carboxy-terminal tail (CTT) helix is displaced by two residues (Fig. 2). This change in sequence register offsets the invariant FFW motif along the helix axis such that Phe 534, and not Trp 536, approaches closest to the flavin ring (Fig. 3). In the new model, the three residues composing the FFW motif continue to make extensive interactions with the photolyase homology domain. This new position of the FFW motif is more consistent with the cellular data of Supplementary Fig. 6 of ref. 1, which shows that substitution of FFW to three alanine residues has a marked effect on dCRY stability, but that the Trp536Ala substitution alone does not. The configuration of the flavin centre is similar between the old and new models, with the largest difference in the angle of the ribityl-to-flavin (N10) bond (Fig. 4). Phosphorylation of Thr518 is not apparent in the new



**Figure 1 | Superposition of Ca traces for 3TVS and rebuilt dCRY structure.** 3TVS is in yellow, and the rebuilt dCRY structure is in cyan. Both subunits within the asymmetric unit are depicted. Flavin cofactors are shown in orange.

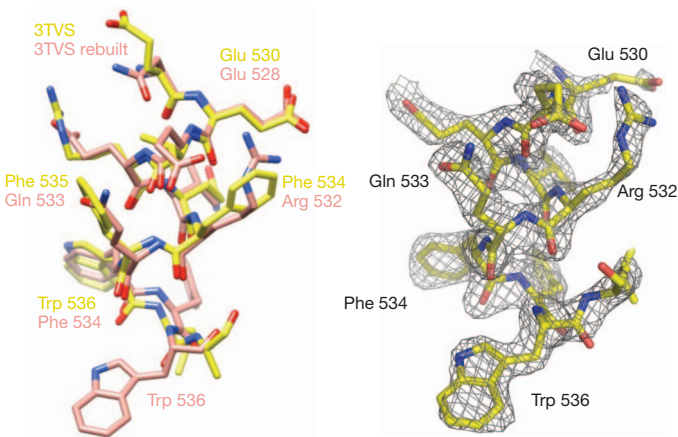
electron density maps despite identification of this modification by mass spectrometry.

The errors in the original structure stemmed from model bias introduced during the detwinning procedure. Lower resolution data sets to which the original dCRY structure was built appear to suffer more from twinning than the 2.3 Å resolution data that the final model was refined against. Although the high-resolution data does contain indications of non-merohedral twinning, including intensity oscillations along the reciprocal space l axis and spurious Patterson peaks, a model that agrees well with the diffraction data as collected can be produced without compensation for these effects (Table 1).

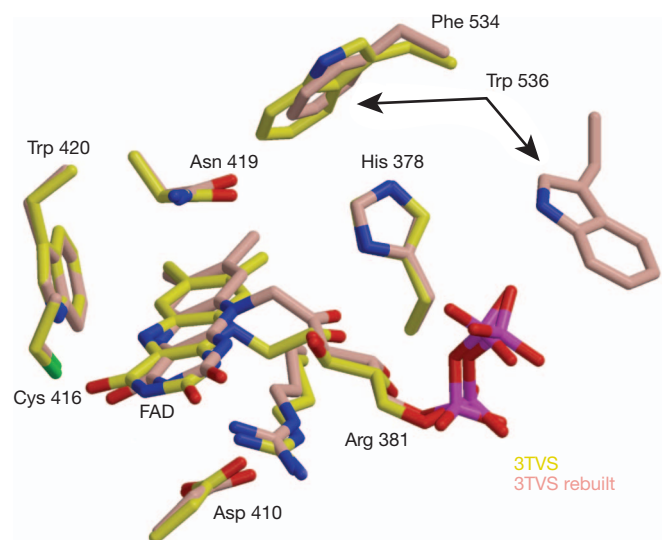
**Table 1 | Data collection and phasing statistics**

	Native	
<b>Data collection</b>	0.97918	
Space group	P2 <sub>1</sub>	
Cell dimensions		
a, b, c (Å)	72.8, 112.3, 75.0	
α, β, γ	90.0, 114.9, 90.0	
Resolution (Å)	30–2.30 (2.38–2.30)*	
R <sub>sym</sub>	0.126 (0.354)	
I/σ (I)	11.9 (1.8)	
Completeness (%)	90.1 (57.3)	
Redundancy	3.9	
<b>Refinement</b>	3TVS (detwinned/native)	Rebuilt (detwinned/native)
Resolution (Å)	30–2.3 (2.30–2.38)*	30–2.3 Å (2.30–2.38)*
Number of reflections	47,507	47,507
R <sub>work</sub> (%)	25.3 (30.0)/34.2 (43.9)	22.2 (24.4)/18.4 (24.2)
R <sub>free</sub> (%)	29.8 (35.8)/40.5 (49.3)	26.0 (34.2)/24.7 (32.1)
Protein	8,668	8,653
Solvent	301	321
Ligand (FAD/Mg)	106/2	106/2
B-factors (Å <sup>2</sup> )		
Protein	38.1	32.2
Ligand (FAD/Mg)	10.0	18.1
Water	28.6	30.7
r.m.s.d.		
Bond lengths (Å)	0.01	0.008
Bond angles (°)	1.8	1.1

Revised data collection and Phenix refinement statistics (corrected Supplementary Table 2 from ref. 1).  
\* Highest resolution bin for compiling statistics.



**Figure 2 | Sequence shift in the CTT.** Left, superposition of CTT from 3TVS (yellow) and rebuilt structure (pink). Sequence register in CTT of rebuilt structure is displaced two residues towards the C terminus compared to 3TVS. Right, electron density for the CTT using phases derived from the new model. The  $F_{\text{obs}} - F_{\text{calc}}$  omit map electron density is contoured at  $3\sigma$ .



**Figure 3 | Flavin centre of dCRY.** The corrected Fig. 3c from ref. 1 is shown. Superposition of key residues in dCRY flavin centre between 3TVS (yellow) and rebuilt structure (pink). Position of Trp 536 and linkage between flavin ring and ribityl chain differs in rebuilt structure compared to 3TVS.

The new atomic coordinates have been deposited in the Protein Data Bank as 4GU5.

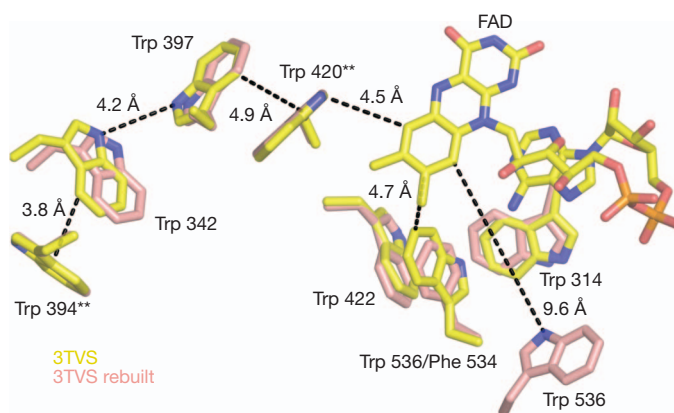
**Colin Levy<sup>1</sup>, Brian D. Zoltowski<sup>2</sup>, Alex R. Jones<sup>1</sup>, Anand T. Vaidya<sup>2</sup>, Deniz Top<sup>3</sup>, Joanne Widom<sup>2</sup>, Michael W. Young<sup>3</sup>, Nigel S. Scrutton<sup>1</sup>, Brian R. Crane<sup>2</sup> & David Leys<sup>1</sup>**

<sup>1</sup>Manchester Institute of Biotechnology, Faculty of Life Sciences, University of Manchester, 131 Princess Street, Manchester M1 7DN, UK.

<sup>2</sup>Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, USA.

<sup>3</sup>Laboratory of Genetics, The Rockefeller University, New York, New York 10065, USA.

e-mail: bc69@cornell.edu



**Figure 4 | dCRY Trp triad photoreduction pathway.** The corrected Fig. 4a from ref. 1 is shown. Arrangement of CRY Trp residues in 3TVS (yellow) and rebuilt structure (pink). Trp 536 is replaced by Phe 534 in the rebuilt structure. Trp 536 is now separated from the flavin ring by the imidazole side chain of His 378 (see Fig. 3). Double asterisks denote that the new positions of Trp 394 and Trp 420 are nearly identical to the old.

Received 12 June 2012; accepted 15 November 2012.

1. Zoltowski, B. D. *et al.* Structure of full-length *Drosophila* cryptochrome. *Nature* **480**, 396–399 (2011).
2. Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. The Phenix refinement framework. *CCP4 Newsllett.* Number 42, contribution 8 (2005).

**Author Contributions** The New York authors (B.D.Z., A.T.V., D.T., J.W., M.W.Y. and B.R.C.) carried out the work presented in ref. 1; the Manchester authors (C.L., A.R.J., N.S.S. and D.L.) identified errors in the original coordinates; together, both sets of authors carried out the further refinement of the structure presented here. C.L. initially identified errors in the original coordinates, and C.L. and D.L. performed initial model rebuilding and crystallographic analysis, which was verified and discussed with A.R.J. and N.S.S. The New York authors B.D.Z. and B.R.C. completed and verified the model rebuilding and crystallographic analysis. The New York authors A.T.V., D.T., J.W. and M.W.Y. contributed to structural analysis and verified the new structure against existing biochemical and biological data. B.R.C. and B.D.Z. wrote the communication with input from all other authors.

**Competing Financial Interests** Declared none.

doi:10.1038/nature11995